

6/167

From: Pak, Michael
 Sent: Wednesday, February 27, 2002 10:15 AM
 To: STIC-Biotech/ChemLib
 Subject: 08/816,011 sequence search

Sequence search - 2 month amendment

App. #: 08/816,011

Result format: Paper.

Title: potassium channels ...

Please search:

Search commercial and interference database.

SEQ ID NO:36 and 63.

Thanks,

Mike Pak

Michael Pak
 Art Unit 1646
 Mailbox: CM1, Rm. 10C00
 Office: CM1, Rm. 10E13
 703-305-7038

Michael Pak
 USPTO
 Art Unit 1646
 CM1; Rm. 10E13
 703-305-7038

*CIP of PCT/US05/14364 10/28/05
 CIP of 08/332,312 10/31/94 Pot-a 5558026
 only 99%
 no support*

RECEIVED
 FEB 27 2002
 (STIC)

if Contact:
Sheppard

Searcher:

Phone: tel: 308-4499

Location: _____

Date Picked Up: _____

Date Completed: 3/1/02

Searcher Prep/Review: _____

Clerical: _____

Online time: _____

TYPE OF SEARCH:

NA Sequences: _____

AA Sequences: _____

Structures: _____

Bibliographic: _____

Litigation: _____

Full text: _____

Patent Family: _____

Other: _____

VENDOR/COST (where applic.)

STN: _____

DIALOG: _____

Questel/Orbit: _____

DRLink: _____

Lexis/Nexis: _____

Sequence Sys.: _____

WWW/Internet: _____

Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

NCBOM protein - protein search, using sw model

Run on: February 27, 2002, 16:59:59 ; Search time 25.85 Seconds
(without alignments)
2455.791 Million cell updates/second

Title: US-08-8116-0111-63
Perfect score: 2294
Sequence: 1-MT-FAE-RNTVAYEQEERFPRD.....LTGGIWMVVIEHFVDKPSIL 434

Scoring table: BLOSUM62
Gapext 0.5
Gapop 10.0

Searched: 473505 seqs, 146273329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

THE SILENT INVISIBLES

```

>Database : SPTRIMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_prok:*
5: sp_euk:*

```

```
sp_urbatelebrate:*
sp_mammal:*
sp_mhc:*
sp_organelle:*
sp_phage:*
sp_plant:*
sp_rrodent:*
sp_virus:*
```

13: sp_vertebrate,*
 14: sp_undefined:*

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2265	98.7	434	5	Q19495	caenorhabdi
2	802	35.0	418	5	Q0U318	caenorhabdi
3	716.5	31.2	461	5	Q23599	caenorhabdi
4	706.5	30.8	450	5	Q93871	caenorhabdi
5	645	28.1	458	5	Q9vmb6	drosophila
6	574.5	25.0	800	5	Q21145	caenorhabdi
7	367.5	16.0	475	11	Q99655	nus musculi
8	343.5	15.0	456	4	O9ujy2	homo sapien
9	336.5	14.7	456	6	Q9B985	oryctolagus
10	328.5	14.3	456	11	Q9J7T8	nus musculi
11	327.5	14.3	475	4	Q9BZD2	homo sapien
12	320.5	14.0	428	10	Q9M5X9	nus musculi
13	313.5	13.7	428	10	Q64603	arabidopsis
14	302.5	13.2	458	11	Q9JF00	nus musculi
15	302.5	13.2	458	11	Q9dB18	nus musculi
16	301.5	13.0	460	11	Q9JTM1	nus musculi
17	301.5	13.1	460	11	Q99R84	nus musculi
18	291.5	12.7	476	5	Q9NHZ1	drosophila
19	261.5	11.4	397	4	Q9NIS9	homo sapien
20	250.5	11.4	397	4	Q9NIS9	homo sapien

Q9Fy1	arabidopsis
Q9VU20	drosophila
Q9VP0	drosophila
Q9GTP5	crittridium f
Q93650	caenorhabdi
Q9Bq4	oryctolagus
Q76343	leishmania
Q76269	leishmania
Q9Bw12	homo sapien
Q9N9R1	leishmania
Q20396	caenorhabdi
Q9nbv4	leishmania
Q9U763	trypanosoma
Q9sr64	arabidopsis
Q9m03	arabidopsis
Q9cvql	toxoplasma
Q9m02	arabidopsis
Q9Y010	trypanosoma
Q9ih8	plasmidium
Q9ckv7	pasteurella
Q9Y09	trypanosoma
Q9GTP4	crittridium f
Q9Bv3	schistosoma
Q99510	trigona ama
Q9nh7	plasmidium
Q9SY87	arabidopsis

ATTACHMENT C

RESULT	1
Q19495	PRELIMINARY;
ID	Q19495
AC	Q19495;
DT	01-NOV-1996 (TREMBLref. 01, Created)
DT	01-NOV-1996 (TREMBLref. 01, Last sequence update)
DT	01-JUN-2001 (TREMBLref. 17, Last annotation update)
DE	SIMILARITY TO HNP36 PROTEIN.
GN	F6H11.3
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditidae.
OC	Rhabditidae; Peloderaiae; Caenorhabditis.
OX	NCBI_TaxID:6239.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BRISTOL N2;
RX	MEDLINE=94150718; PubMed=7906398;
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA	Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA	Gardiner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
RA	Jones M., Kershaw J., Kirstein J., Laird N., Latrellie P.,
RA	Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
RA	Parsons J., Percy C., Riffen L., Roopa A., Saunders D., Showman R.,
RA	Smalldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA	Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA	Watson A., Weinstock L., Wilkinson-Spiro J., Wohldman P.,
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans."

Query Match	31.28;	Score 716.5;	DB 5;	Length 461;		DR ProDom; PD005103; DER_Nucleoside_tran; 1.
Best Local Similarity	35.88;	Pred. No. 1.2e-39;			SQ SEQUENCE	450 AA; 50495 MW; CF397D5EE517909 CRC44;
Matches	163;	Conservative	86;	Mismatches	167;	Indels 39; Gaps 9;
Matches	11	NKTKVKEESEPPEDKGNLVFLYILLHGIGTIMPWNMLITISDYFESYKMLANSTIDMD 70			Query Match	30.8%; Score 706.5; DB 5; Length 450;
Qy	7	SNTYABQEAA-PPRDKVNIVWVLVILVGFGILPWNMFITAPEYYVNVWWKPD-----	59		Best Local Similarity	36.0%; Pred. No. 5.5e-39;
Db	11				Matches	157; Conservative 88; Mismatches 164; Indels 27; Gaps 8;
Qy	60	-----GYETWYSKEFMGSLTIGSOLNPASINVFNLFLIAGPLIYRVFAPCFNIVNLT 114			Qy	12 VEQEA-FPRDKVNIVWVLVILVGFGILPWNMFITAPEYYVNVWF-----KPD 59
Db	71	TGVTGTYPTVYSSNFQSQTIAQVNPNLNLNIFIVKGGLASR1-TVGLSIVAVCV 128			Db	21 VEEETPEPEDKGNLVFLYILLHGIGTIMPWNMLITISDYFESYKMLANSTIDMDTGKV 80
Qy	115	ILJLIVILEPTEDSMSKRFWWTLMGATNSINFNSGLYENSYVGGGDEPHTYIGALLIGNN 174			Qy	60 GWTWYSKEFMGSLTIGSOLNPASINVFNLFLIAGPLIYRVFAPCFNIVNLTILILLY 119
Db	129	ITMMFYIYVETSTWLWTFPFLTITIVLNGANGYVONSICLASELPFKTNAVIGN 188			Db	81 GDETVVYSSNEFOSQTIAQVNPNLNLNIFIVKGGLASR1-TVGLSIVAVCVTTMI 138
Qy	175	ICLLITVVVKI---GVTVFLNDEPKLVIAVYFGISLVLLCAIALEFFITKODFYHYHQQ 231			Qy	120 IVLEPTEDSMSWFWVTLGMATSINFSNLGYLENSVYVGDDPHTYIGALLIGNNIGGL 179
Db	189	LGCTFTYLMSMSTKAATRNILDR---SEAFASTALITLVCFCISHLKQRFYQYST 244			Db	139 FIVYETSTWLGFETITITIVLNGANGYVONSICLASELPFKTNAVIGNNLGTF 198
Qy	232	KGMEIREK--AETDRDPSPSILWTTFCYGOLENWFCFAVLTLLIP-VMATVTRGDS 287			Qy	180 ITVVKI---GVTVFLNDEPKLVIAVYFGISLVLLCAIALEFFITKODFYHYHQQMEI 236
Db	245	RABQRQNKKNDAEVDESGKVNIAITKEAFLPLINNLFVFTLSIFPGVMYVKDEKKG 304			Db	199 VTLLSMSTKAATPNILDR---SFAYSSIALITLVEFISFHILKKRFYQFSTRERQ 254
Qy	288	G-----FLANKISENDEI---YTLTSLEVNLFAAIGSIVASKTHWHPTRYLK 333			Qy	237 R---EKAEETDRDPSPSILWTTFCYGGLENWFCFAVLTLLIP-VMTVTTRGDSGPFLNK 292
Db	305	GYDFPLPROFFINSLETHNREFSEKYFMDTITFLQVNVFVFIGTSIVAKRQWPAPKLW 364			Db	255 RAKNEEADNEGRMANIYATEKFABPOLINPVLVEFTLSPGVMYVVKDEKKGGYDF 314
Qy	334	FATILRALFIPPFCCNYRVQTRAYPVFFESTDIFVGGIMAFSFSHCGYLSALAMGYTPNV 393			Qy	293 IMSENDEIYLTTSFLYFNFLFAIGSIVASKTHWHPTRYLKFAITLIPFIPFFFCNYR 352
Db	365	IPVYLRLIYPPIFCCNYLPETRSLPFFESTWLFLVIAAMSFGSGYFSGLAMMTTSKT 424			Db	315 PLPQN- -YFMDVTTFLQFNVFATGSIYVAGRQWPAPKLWIPVYLRLYEFFFCNYL 372
Qy	394	VPHYSRSFQAQLSVCTLMVGLTGGIMPVVIHFV 428			Qy	353 VQTRAYPVFFESTDIFVGGIMAFSFSHCGYLSALAMGTTPNVPVPSHYSRFAAQLSVCTLMV 412
Db	425	VDFSKAQVAGMAGFFLISGIVSGLIFTMVKMWV 459			Db	373 PETRTLPPFFESTWLFLVIAAMSFGSGYFSGLSMMYTSKTVDPSSKAVQVAGMAGFFPLIS 432
RESULT	4				Qy	413 GLUTGGLWPMVVIEHFV 428
ID	Q93871	PRELIMINARY;	PRT;	450 AA.	Db	433 GLIVSGLITTMVAKFV 448
AC	Q93871;					
DT	01-FEB-1997	(TREMBLrel. 02, Created)			RESULT	5
DT	01-FEB-1997	(TREMBLrel. 02, Last sequence update)			Q9WMB6	PRELIMINARY;
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			ID	Q9WMB6;
DE	K09A9.3	PROTEIN.			AC	Q9WMB6;
GN	K09A9.3				DT	01-MAY-2000 (TREMBLrel. 13, Created)
CA	Caenorhabditis elegans				DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditida; Rhabditidae; Peloderaiae; Caenorhabditis.				DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
OC	NCBI_TAXID=6239;				DE	CG11045. PROTEIN.
OS	Drosophila melanogaster (Fruit fly).				GN	BEST: LDO971 OR CG11045.
OS	Eukaryota; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidea; Drosophilidae; Drosophilidae; Drosophilidae.				OS	Drosophila melanogaster (Fruit fly).
NCBI_TAXID=7227;					OC	OC
RN					NCBI_TAXID=7227;	NCBI_TAXID=7227;
RP	SEQUENCE FROM N.A.				RN	[1]
RP	SEQUENCE FROM N.A.				RP	SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.				RC	SEQUENCE FROM N.A.
RX	MEDLINE=94150718; PubMed=7906398;				RX	SEQUENCE FROM N.A.
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Craxton M., Dear S., Durbin R., Favell A., Fulton L., Gardner A., Green P., Hawkins T., Johnston L., Jones M., Kershaw J., Laird N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Riffen L., Rooper A., Saunders D., Shownkeen R., Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaughan M., Watsonson R., Watson A., Weinstock L., Wilkinson-Spratt J., Wohldman P.;				RA	SEQUENCE FROM N.A.
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans"; EMBL: 368:32-38(1994); DR: 279601; CAB01882.1; InterPro: IPR00259; DR: Nucleoside_tran; Pfam: PF01733; Nucleoside_tran; 1.				RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Cherry J.M., Cawley S., Dahl C., Davies P., de Pablos B., Delcher A., Deng I., Dietz S.M., Dodson K., Douq L.E., Downes M., Dugan-Rocha S., Dunk P.,

RESULT	6	Q99P65	7		
Q21145		ID	Q99P65		
TD	Q21145	PRELIMINARY;	PRELIMINARY;		
AC	Q21145;	AC	Q99P65;		
DT	01-NOV-1996 (TREMBrel. 01, Created)	DT	01-JUN-2001 (TREMBrel. 17, Created)		
DT	01-NOV-1996 (TREMBrel. 01, Last sequence update)	DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)		
DT	01-JUN-2001 (TREMBrel. 17, Last annotation update)	DT	01-JUN-2001 (TREMBrel. 17, Last annotation update)		
DE	KO2E11.1 PROTEIN.				
RA	KO2E11.1.				
RA	Caenorhabditis elegans	OS			
RA	Eukaryota; Metazoa; Nematoidea; Rhabditida; Rhabditidae;	OC			
RA	Peloderaida; Peloderinae; Caenorhabditis.	OC			
RA	[1]	NCBI_TaxID:6239;	OX		
RA	RN	RN			
RA	RP	RP			
RA	SEQUENCE FROM N.A.				
RA	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.	RL			
RA	SEQUENCE FROM N.A.	RN			
RA	SEQUENCE FROM N.A.	RP			
RA	MEDLINE-R-94150718; PubMed=7906398;	RX			
RA	WILSON R., Ainscough R., Anderson K., Baynes C., Berks M., Cooper J., Coulson A., Bonfield J., Burton J., Connell M., Copsey T., Fultton L., Gardner A., Dear S., Durbin R., Favell A., Fulton L., Green P., Hawkin T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten T., Laird N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Roopra A., Saunders D., Shownkeen R., Simlald N., Smith A., Sonnhammer E., Staden R., Sulston J., RA	RA			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirska S., Tector C., Turner E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zheng L., Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., RT	RA			
RA	"The genome sequence of <i>Drosophila melanogaster</i> ," Science 287:2185-2195 (2000).	RA			
DR	EMBL: AE003614; AACF12405; 1.	DR			
DR	FlyBase: FBgn0026585; BEST; DR	DR			
DR	InterPro: IPR002259; DER_Nucleoside_tran.	DR			
DR	PF01733; Nucleoside_tran; 1.	DR			
DR	PRINTS: PRO01130; DERENTRNSPRT	DR			
DR	Prodrom; PD005103; DER_Nucleoside_tran; 1.	DR			
SEQUENCE	458 AA; 51324 MW; 72D36D4ICE0D4A2E CRC64;	SEQUENCE	800 AA; 90382 MW; E056D51129EE4BF45 CRC64;		
Query Match	28.1%	Score 645; DB 5; Length 458;	Query Match	25.0%	Score 574.5; DB 5; Length 800;
Best Local Similarity	32.9%	Pred. No. 6.1e-35;	Best Local Similarity	32.7%	Pred. No. 4.4e-30;
Matches	138; Conservative	Mismatches 178; Indels 20; Gaps 7;	Matches	134; Conservative	Mismatches 165; Indels 23; Gaps 10;
QY	18 PRDKYNIVWLVILVGFGVLLPWNNMFITTAPEYYVNYWFKPD---GVETWYSKEFMGSILT 74	QY	19 RDKYNNIVWLVILVGFGVLLPWNNMFITTAPEYYVNYWFKPDGVETWYSKEFMGSILTGS 77		
Db	51 PKDFPLIVFFFLHGVTLMWNMFITAKSYEDFKFGPNNTVATESYRTHFMQNMG 109	Db	388 KDGNYVFTIFMAGFGGALLPWNNMFLNISFDYUTMFKLRSADGNATWSSNFNSMTISA 447		
QY	75 IGSOLPNASTINVNLFILITAGPLIYRFAPCFNIVNLTLVILEPTEDSMSW-- 131	QY	78 QLPNASIWINVNLFILITAGPLIYRFAPCFNIVNLTLVILEPTEDSMSWFWTIL 137		
Db	110 FASQPLNVLVNLFVNFGDLTRVISITFEMVLLVITLAML---DSSQWPGV 164	Db	448 QIPSLVFSVINITAVKGDLTRGM---KICLIVVQLMVTVVVFYIDSTWIATEFSMLT 505		
QY	132 FEWTLGATMSINFNLGTSVSYVGCGDEPHTYIGLGNNCGLLTIVKVGTFLNDEEFL 197	QY	138 GMATSFINGLVSFNGVSYVGCGDEPHTYIGLGNNCGLLTIVKVGTFLNDEEFL 197		
Db	165 FEWTTMVCIVLNVNGTQNTIGTIVASLPIKTYGAVVLGSNISGCTTAMAL-ICBEI 223	Db	506 GTIVVLNAANGLFRNSMGLASPPFKTNAVLQNCFCGTAIVVLSM-LTKASDDQJM 564		
QY	192 NDEPKLVAIVFGSLVLLVCAATALFFITKQDFHYHHQKGMEIREKAETDRPSPLWL 251	QY	198 VATVYFGISIVLVLVCAATALFFITKQDFHYHHQKGMEIREKAETDRPSPLWLTT---- 253		
Db	224 FSSKRTSAIYYFTAILVLLCFTOTYFALPLNKFRERHETISSESEKKSDSRQLNVYW 283	Db	565 RASLFFGGLSSVAVVCFILLNFKRLAY---KKFGFLRTSSQSDEGGIS-SWEVSLA 619		
QY	252 TTFNCGYOLFNYWFCFAVLTIEPVMTVTGDSGLINKSENDEIYTLLTSLFLVFN 311	QY	254 FTNCYQOLFNPWRCFAVLTIEPVMTVTGDSGLINKSENDEIYTLLTSLFLVNLF 313		
Db	284 QIEKKAAPQFLNIELTFFVTLVEPAIQSNVRSDPNEV---VGPDYFTLYTCFATEN 338	Db	620 FEKSKMQANIEVLEFFVTLAFPNCVMVKDAKGELISFVVP-EKYMDDVVTFLNFILF 678		
QY	312 LFAATGSIATSKHWPYKFAITRPLKFAITRPLKFAITRPLKFAITRPLKFAITRPL 369	QY	314 AATGIVASKHWPYKFAITRPLKFAITRPLKFAITRPLKFAITRPLKFAITRPL 371		
Db	339 VFAMIGSLTSITSWQWPGRFLWVPPVLRALAFPLFVPMCNYPDSVRSVLAVFTENDWWYW 398	Db	679 AFLGSLMANWIRFPGPNTWICVAARFWFMFYPAANNHPMDPRAVEPPVLFESTWLFFAFN 738		
QY	370 IGGTAMSFSHGYLSLALANGTYPNVPSPHSYRSAQLSVCTLMYGLLTGGWMWVIEHFV 428	QY	372 GIAMSFSHGYLSLALANGTYPNVPSPHSYRSAQLSVCTLMYGLLTGGWMWVIEHFV 417		
Db	399 GITAMAYSSGGLSLGMMYAPQTAGMYAAAMLITGIFSGVLFYLGPFRRV 457	Db	739 ICIFALTSGYLLSLIMMYAPR---SHEDPKIQRMAGMMASFFLIFGTVAG 785		

DE	EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 3.	RT	"Genomic Sequence of the human Equilibrative Nucleoside Transporter 1 (ENT1)." ;
GN	MUS MUSCULUS (Mouse).	RT	Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RL	[2]
OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus;	RN	SEQUENCE FROM N.A.
OX	NCBI_TAXID=10090;	RC	TISSUE-COLON ADENOCARCINOMA;
RN	[1]	RA	Strausberg R. ;
RP	SEQUENCE FROM N.A.	RA	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RC	STRAIN=C57BL;	DR	EMBL; AE190884; AAF02777.1; -;
RA	Hyde R.J., Cass C.E., Young J.D., Baldwin S.A. ;	DR	EMBL; BC001982; AAH01382.1; -;
RT	"The ENT family of eukaryotic nucleoside and nucleobase transporters: recent advances in the investigation of structure/function relationships and the identification of novel isoforms." ;	DR	InterPro; IPR002459; DBR_Nucleotide_tran.
RT	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ	DR	PRINTS; PF01733; Nucleoside_tran; 1.
DR	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ	DR	PRODOM; PD005103; DER_Nucleoside_tran; 1.
DR	EMBL; AF326386; AAK00957.1; -;	DR	SEQUENCE 456 AA; F1855A95DB95D CRC64;
SQ	SEQUENCE 475 AA; 51719 MW; F59315D03C6CBB52 CRC64;	DR	50219 MW; F1855A95DB95D CRC64;
Query Match	15.0%	Score	343.5; DB 4; Length 456;
Best Local Similarity	25.8%	Prd. No.	3..5e-15; Gaps 13;
Matches	123; Conservative	Mismatches	77; Indels 59; Gaps
Qy	7 SNTYAVEQEAF-----PRDKYNIVVWLYLVFGVLLPWNMFTIAPEYYV	Qy	18 PRDKYNIVVWLYLVFGVLLPWNMFTIAPEYYVWNKFDPGV-----TWSKEFNGSLT 74
Db	21 SNHQEADQEAJLGKLIDYPAPGLQRDFEDRFNGAYIIFCLGIGGLLPWNFPVT-AKEYWA	Db	7 PQDRYKAWLUFFMLGIGTLPWNFTW-ATQYETNRDMSQNLYVTAELSMDAQASAA 65
Qy	53 NYWFKPQDGVEWTYSKE-----FMGSLTIGSOLPNASAINVNFLITAGPLIYRVF	Qy	75 TGSQLP-NASINVNFNLITAGPLIYRVF-----PVCFNIVN-----LTILII 117
Db	80 ---YKLRNCSSPASGEDPEDMILNFEYSYLAVALSPVSLFLVANFLVNRVQHVRL	Db	66 PAAPPLPRNSLSAIFNNVMTLCAMLPLLFYIILNFSWFFVTLGMATSIINSFNSGLYENSYVGCGDFPHTYIGALLIGNNCG 177
Qy	103 APVCFNIVNLTTILIVLIVLEPTDSMW----FWWYTLGMATSINSFNSGLYENSYVGCG 159	Qy	118 LTVLEPTEDSMNSWFFVTLGMATSIINSFNSGLYENSYVGCGDFPHTYIGALLIGNNCG 177
Db	137 ASISVSUAIYFVINYLKV----DTSWTFSEUTIACHAIISSSTTNSYGLTG	Db	126 TAILVKYQLDALP-FFVITMIKIVLINSFGAILQLOGSFLGAGLPPASYTAPIMSGOGLAG 184
Qy	160 DEPHTYIGALLIGNNIGLITVVKIGTYTFINDEPKLVAVYFGISLVLVCAIALFF	Qy	178 LLITVVVKIGTYVFLNDEPKLYAIVYFGISLVLVCAIALFFIKDQFYHFO-----231
Db	192 SPPMRNQAQLSGGAMCG-TYSAVALLVLDLASSDVDRDSTLAFFELMAAVFLGLCMGLYLL	Db	185 EFASVANI-CAIASGSLSESASFAGYFYTACAVILLTICYLGLPRLFIFYQQLKLEGP 243
Qy	220 TTKQDFVHYHHQKGME TREKAETDRS---PS-----ILWTFETNCY	Qy	232 -----KGMETREKAETDRPSISLWTT-----FTNCGQLENWNPCEFAV 271
Db	251 LSOLEYARYMPVAPVPRVFGEDNNSQDASSVAPASRVMMHTPPLGPPLKKT-----	Db	244 GEQETKDLISKGEPPAGKEESGVTSNSQFTNEHSIKIALLKNISVLAFCVCFIT 303
Qy	259 GOLFNYWFC----FAVLTIPVMFTVTRGDGFLINKIMSENDEIYTLITSEFLYENLFA	Qy	272 LTFPVMMTVTRGDGFLINKIMSENDEIYTLITSEFLYENLFAIGSIVASKIHW-----TP 329
Db	306 ---ASLGFCAVSLYFNTAFAPIASTNIQSHKGTSPPWTS---KEFVPLVFLLENFAD	Db	304 IGMPATVVEKSSIAHG-----SSTWERYFIPVSCFUTENFDWLGRSLTAVFMWPGKDS 358
Qy	315 AIGSVASKIHWPTPR--YIKFAILRLAFLPFPEFFCNRYQOTRAYPFFESTDIF-VIG	Qy	330 RYLKFAITLRAFLPFPEFFCNRYQOTRAYPFFESTDIF-YIG-----388
Db	360 LGCGROYTAIWQVPGRSKLPLQVSRFLCNQYQRSHLTKVLFQS-DIYPVLF	Db	359 RWLPSLVLARLYEVPLLLCN--IKPRYLTWVFEHDAWTFMMAAFNSGNYLASICMC 416
Qy	372 GIAMSFSHGYSALAMGYTPNVPVSHYSRFAAQLSVCTLMGLITGGLWPVIEHHFV	Qy	389 YTPNVVPSHYSRFAAQLSVCTLMGLITGGLWPVIEHHFV 428
Db	419 TCGLGSNGYLSTLVLIYGPKVPRELAEEASVVMFLYMSVGLMLSACAALEHFI	Db	417 FSPKKVVRPAEATAAGMAFFLCLGLPLGAVALSFELPRAIV 456
		RESULT	9
Q9UJY2	PRELIMINARY;	PRT;	456 AA.
ID	Q9BG85	PRELIMINARY;	PRT;
AC	Q9UJY2	PRELIMINARY;	PRT;
DT	01-MAY-2000 (TREMBLrel. 13, Created)	AC	Q9BG85;
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 1 (SOLUTE CARRIER FAMILY 29 (NUCLEOSIDE TRANSPORTERS), MEMBER 1).	DE	NBMPR-1-SENSITIVE NUCLEOSIDE TRANSPORTER EI.
GN	ENT1.	GN	ENT2.
OS	Homo sapiens (Human).	OS	Oryctolagus cuniculus (Rabbit).
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Leporidae; Lagomorpha; Leporidae; Oryctolagus.	OC	NCBI_TaxID=9986;
OX	[1]	RN	SEQUENCE FROM N.A.
RN	RP	RN	Wu S.K., Ann D.K., Lee W.H.L. ;
RA	Sequence FROM N.A.	RA	"Molecular cloning and functional characterization of equilibrative nucleoside transporters (rbENT2 and its constitutively spliced variant rbENT2A) in rabbit."
RA	Graham K.A., Coe I.R., Carpenter P., Baldwin S.A., Young J.D.,	RT	
RA	Cass C.E.,	RT	

RL Submitted (NOV-2000) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AF323951; AAK11605_1; .
SQ SEQUENCE 456 AA; 49858 MW; C6D55FC9DD2A392 CRC64;

Query Match 14.7%; Score 336.5; DB 6; Length 456;
Best Local Similarity 24.9%; Pred. No. 1e-14;
Matches 115; Conservative 68; Mismatches 191; Indels 95; Gaps 14;

Qy 18 PRDKYNIVYWLVILVGFGVLLPWNMFITTAPEYYVNNYWFQKDPGEVETWYSKEFNGS---- 72
Db 7 PRDSYHLVGISFFFLGLTLLPWNMFITTAIP---- YFQGRLAGTNNSAETGNTNHS 60

Qy 73 -----LTIGSQLPNAISINVNLFLTAGPLIYRFAPIVCFNIVNUTILLVILE 123
Db 61 TDTEFNWNVTLISQLPPLFTLNSFLQCIPESRVRL----- SLLALLFALTAALV 116

Qy 74 PTEDSMSNEFWYTIGMATSINFNSGLYENSYVGGDFPHTYIGALLGNICGLLTVV 183
Db 62 117 KVDLSPGIFFSVNTMSAVNFSQAVQSLFGQLGPSTVSLFGQGLAGIFALA 176

Qy 124 117 KVDTSPGIFFSVNTMSAVNFSQAVQSLFGQLGPSTVSLFGQGLAGIFALA 176
Db 117 KVDLSPGIFFSVNTMSAVNFSQAVQSLFGQLGPSTVSLFGQGLAGIFALA 176

Qy 184 KIGTYFLNDEPKLYAIVF----GISLVLLCAIALFFITKQDFYHFFHOK----GM 234
Db 182 ASGV-----DAQTSALGYFLTPCGIVLIVSVCYLSPHL----- 230

Qy 221 -----TKQDFYHYHKGMEIR-----EKAEDRSPSPSILWTTFTNCYG 259
Db 231 TQELETKAELLQADEKNGNPISPOQASSTLDLDEKEPEEEQKPGPSPVFVRKJWL 290

Qy 260 QLNWVWFCEAVTLLTIEPM----MTVYTRGDSG----FLKIMSENDEIYLLTSLFLVNLF 313
Db 291 TAICLVLYFLTVTUVSPALTAMYNTSSNPGKNGLFFPVI-----CCFLFLFVNU 339

Qy 314 AAIGSIVASKIHW-----TPRYLKFAITIRALFPEFFCNYVQTRAYVPPFESTDIVI 370
Db 340 DWLGRSLTSYFLWPDEDQOLLELLVCLRFLEPFMLCHPVQHAR-LPIIIFRDAYIT 398

Qy 371 GGIAAMSFGHGSALAMGTPTNVYPSHYSRFAAQLSYCTLMVGLLTG 417
Db 399 FMLLFAVSNGLVLSLTMCAPRQLPHREVAQALMTFFLALGLSCG 445

RESULT 11
Q9BZD2 PRELIMINARY; PRT: 475 AA.
ID Q9BZD2; PRELIMINARY;
AC Q9BZD2; PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 3.
GN EN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Hyde R.J., Cass C.E., Young J.D., Baldwin S.A.;
OC "The ENT family of eukaryotic nucleoside and nucleobase transporters:
RT recent advances in the investigation of structure/function
RT relationships and the identification of novel isoforms.";
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RL DR EMBL: AP326937; ARK00938.1; .
SQ SEQUENCE 475 AA; 51874 MW;

Query Match 14.3%; Score 327.5; DB 4; Length 475;
Best Local Similarity 24.4%; Pred. No. 4.e-14;
Matches 116; Conservative 83; Mismatches 202; Indels 75; Gaps 15;

Qy 7 SNTYAYEBOAF-----PRDKYNIVYWLVILVGFGVLLPWNMFITTAPEYYV 52
Db 21 SSSURADOEALLEKLDRPPGLQRPERFRCGTYIIFPSLGIGSLLPWNMFIT-AKEYWM 79

Qy 53 NYWEK-----PDGVETWYSKEFNGSITIGSOLPNASINVNLFLIAGPLIYR 100
Db 80 ---EKLRNSSLSSPATGEDDEGS -ILNWFESTLAVASTVPSMCLVANFLYRNVAHVIR 134

QY	78	QLPNASINVNFLILIA----GPIIYRVFAPVCNFIVNLTLILIVLIVLEPTEDSMSWF	132	DR	PRODOM; PD005103; DER_Nucleoside_tran; 1.
Db	86	-----LVALVCLVIVVYAAHKSLASFRINGLLFLVIALLVPVLDVYRGQVGYAG	140	SQ	SEQUENCE 458 AA; 50076 MW; EDE0EDEF5BF3E274 CRC64;
QY	133	FWYTLGMATSINFNSGNYENSYGVGGDFPHTYIGALLIGNICGLLITVKGTVTFLN	192	Query	Match 13.2%; Score 302.5; DB 11; Length 458;
Db	141	FDTSAAVALSGLGDAALMGGGLVGEMPERYMQAVVAGTAGSAGVLSLLRILTAKAVYP	200	Best Local Similarity 22.5%; Pred. No. 1.7e-12;	
QY	193	DEP---KLVAYVFGISVLLVCAIALEFTTQDFHYHHQKGME---IREVAE-----T	242	Matches 104; Conservative 82; Mismatches 215; Indels 61; Gaps 13;	
Db	201	QDPDGLRKSANLYFAVGIVVMTCAVYNAHLPVKPFHEERKNEELIRESEEEKGSLT	260	Oy	18 PRDKYNIVYWLVLVGFVLLPNNMFITIAPEYYVNVW----FKPDDGVTETWSKEFMGS
QY	243	DRPSPSIWLWTFNFCYGOLENWFCFAVLTIFPVMMTVTTRGDGFLINKIMSENDELYT	302	72	72 PRDKYNIVYWLVLVGFVLLPNNMFITIAPEYYVNVW----FKPDDGVTETWSKEFMGS
Db	261	GLAWRTTLNDIVTKVKSHFGFIVLYMVTLISLFPGYIT-----EDVHS	303	Db	7 PQDRYKAWLFLFVGLGLTLPNNFEM--ATKFTNRDLDVSNQVSSDTQSCSTKAAD
QY	303	-LITSF-----LIVENLFAAIGSTVASKIHWPTPRYLKFIAILRLALPFEFFCNYRVQT	355	Qy	73 LTIGSOLFNASINVENLFLIAGPLIVYFAPVCEN-----IVNLTLIL 116
Db	304	ELTDWYPLLAYNVDLVLGCRCLTAFLVMLDEKIAVGSSARLFLYPLFWGLCLHG---	360	Db	66 PTVALPASSLSAIENVNNTMTCAMLPJLVE--TCLNSFLHORISOSVRILGSLLAIVVF
QY	356	RAPVFFESTDITVIGGIAMFSHGYLSALANGYTPNVPSPHSYRSFRAAQSLCUTLWGLL	415	Qy	117 ILVIVLETEDMSWEFWFLTGMAINSGLYENSYVGVGDFEPHYIGALLIGNIC
Db	361	-- PMFLTEIPVTLTCLGLTNGYLTSVLMLAPSKVPLRSETAGIVTMFLVYGLA	417	Db	124 LVTAALVYKVEMDALIFVITMIKIVLINSFGAILQASFGLAGVLPANYTAPMSGQGLA
QY	416	TGGL--WPVVI 4.24		Oy	177 GLLTVVYKIGTYFLNDEPKLYAIVFGISLVLVCAIALFFITKODFYHYHOKGE-
Db	418	SGSVIAWFWVI 4.28		Db	235 RFLKFAILRLAFLPFFFCNYRVQ--TRAYPVFFESTDIEVGGIAMSFSHGYLSALA
RESULT	14	PRELIMINARY;	458 AA.	Qy	386 330 RYLKFAILRLAFLPFFFCNYRVQ--TRAYPVFFESTDIEVGGIAMSFSHGYLSALA
Q9JHF0	1D	Q9JHF0	AC	Db	357 RWLPVLVSRVTFIPLMLCNVYKARHGQAQRHFEVHDWAFATMFAFAFSNGYLASLC
Q9JHF0	01-OCT-2000	(TREMBrel. 15, Created)		Qy	416 387 MGTPNVPSPHSYRSFRAAQSLCYTLMGLTGGMPVYVIEHFV
Q9JHF0	01-OCT-2000	(TREMBrel. 15, Last sequence update)		Db	417 MCGPKKVKPAAEATAGNIMSFPLCLGIALGAVLSFLRALV
Q9JHF0	01-JUN-2001	(TREMBrel. 17, Last annotation update)		Qy	458 236 -----IRE-----KAEDTRPSP-----SILWTTFTNCYQOLENWWFCFAVT
Q9JHF0	DE	EQUILIBRATIVE NITROBENZYLTHIOINOSINE-SENSITIVE NUCLEOSIDE TRANSPORTER		Db	243 PAEQETKLDLIREPKGREREESGPVGPNSPPPTNRNQSKAIILKSIICVPAL-SVCFITVT
Q9JHF0	DE	ENT1.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
Q9JHF0	GN	SLC29A1 OR ENT1.		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
OC	OC	MUS MUSCULUS (Mouse).		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
NCBI_TaxID=10090;	NCBI_TaxID=10090;	Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
[1]	[1]	SEQUENCE FROM N.A.		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RN	RN	SEQUENCE FROM N.A.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RP	RP	STRAIN=CD1; TISSUE=WHOLE BRAIN;		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RC	RC	Hammond J.R., Kiss A.J.;		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RA	RA	"Molecular cloning of the full-length cDNAs encoding the		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RN	RN	RT nitrobenzylthioinosine-sensitive and nitrobenzylthiinosine-resistant		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RT	RT	RT equilibrative nucleoside transporters from mouse brain."		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RT	RT	RT submitted (FEB-1999) to the EMBL/GenBank/DBBJ databases.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RN	RN	SEQUENCE FROM N.A.		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RP	RP	SEQUENCE FROM N.A.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RC	RC	Choi D., Handa M., Young H.M., Messing R.O.;		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RT	RT	RT "Genomic organization and expression of mouse equilibrative nucleoside transporter 1 (es-ENT1) gene."		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RT	RT	RT Submitted (AUG-2000) to the EMBL/GenBank/DBBJ databases.		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RN	RN	SEQUENCE FROM N.A.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RN	RN	STRAIN=129/SVJ1.		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RC	RC	RC Choi D.-S., Messing R.O.;		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RT	RT	RT "Alternative splicing variants of equilibrative nucleoside transporter 1."		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
RT	RT	RT Submitted (AUG-2000) to the EMBL/GenBank/DBBJ databases.		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
RL	RL	RL EMBL: AF131212; AAF78452.1; .		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
DR	DR	DR EMBL: AF218255; AAF64036.2; .		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
DR	DR	DR MGII; MGII-192/073; SIC29a1. .		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
DR	DR	DR InterPro: IP0002259; DER_Nucleoside_tran. .		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
DR	DR	DR Pfam: PF01733; Nucleoside_tran. .		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS
DR	DR	DR Nordone P., Ring B., Ringwald M., Mombaerts P., .		Qy	329 272 LTIFPPVMNTVTRGDGFLNKIMSENDELYTLLSFVNLFAIGSTSVAKSTHWP--TP
DR	DR	DR Sasaki H., Sato K., Schoenbach C., Saya T., Storch K.-E., .		Db	356 302 IGLFP--AVTAEVESSTAG--TSPWKSYFPIVACLFNLNFVWDWLGRSLTAVCMWPQDS

RA Suzuki H., Toyo-o-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).

RL EMBL; AK004756; BAB2537_1;

DR MGD: MGI:1923527; 120001421Rik.

DR InterPro; IPR00259; DER.Nucleoside_tran.

DR Pfam; PF01133; Nucleoside_tran; 1.

DR PRINTS; PRO0130; DERENTRNNSPRT.

DR Prodrom; PDO0510; DER.Nucleoside_tran; 1.

DR Sequence 458 AA; 50018 MW; 01F300DEBBE4FA0C6 CRC64;

Query Match 13.2%; Score 302.5; DB 11; Length 458;

Best Local Similarity 22.5%; Pred. No. 1.7e-12;

Matches 104; Conservative 82; Mismatches 215; Indels 61; Gaps 13;

Db 7 PDRYKAVWLFVFLGQTLIPWNMFTTIAPEYYVNYW----FKPDGVETWYSKEFMGS 72

Db 117 LTIGSQLPNASINVNFLLITAGPLIYRVPVCFN-----IVNLTIL 116

Db 66 PTVALPARSSLSAIFNNVMTLCAMPLVLFVLPWFMT-ATKYFTNRDVSQNSDTGOSCESTKALAD 65

Db 124 LYTAALVKVENDALIFFVITMIKVILINSFCAILQSLFGIAGVLPAHYQGLA 183

Db 177 GLLITVVKIGTYFLNDPEKLVAVIYFGISIVLIVCATIALEFFTKDGFYHHHQKGME- 235

Db 184 GFEFTYAMI-CIAIASSESSAAGFETTAGVIALYLALPRTEFYHLYLNLAG 242

Db 236 -----IRE-----KETDRPSP-----SILWTFITNCYGOLENWWFCFAVT 271

Db 243 PAEQETKLDLIREEPKGRREESGVGPNSPPTNRNQSIKALLKSCIVPAL-SVCFIFTVT 301

Db 272 LTIFPVMMTVTRGDGFLNIMSENEDIVTLLTSLPENFAAIGSIVASKIHP--TP 329

Db 302 IGLFP---AVPAEVESSIAAG-TSPMKSYFIPVACFLNFVNDWLGRSLTAVCMPPGQDS 356

Db 330 RYLFKAIIIRAFIPIFFFCHYRYQ--TRAYPVPEFESTOFLVIGGTAMSFSHGYSALA 386

Db 357 RWLPLVLAISRIVFIPMLCNVKARICGAQHHEVFKHDAMFLAPMAAFASNGYIASLC 416

Qy 387 MGYTPNVVPSHYSRFAAQQLSUTLMVGLTGGLWPVVIIEFY 428

Db 417 MCFGPKVKPAEAETAGNIMSSFFLCLGLAUGAVLSPFLRALV 458

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2002, 23:22:38 ; Search time 1347.44 Seconds
(without alignments)
16993 .770 Million cell updates/sec

Title: US-08-816-011f-36
Perfect score: 1388
Sequence: 1 acggtaataatccacgatc ttattaaaaaaa 1388

Scoring table: IDENTITY NUC Gapext 1.0
Gap 10.0 , gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenBml:

1: qb_ba:
2: qb_htc:
3: qb_in:
4: qb_cm:
5: qb_ex:
6: qb_pat:
7: qb_ph:
8: qb_pl:
9: qb_pr:
10: qb_ro:
11: qb_sts:
12: qb_sy:
13: qb_un:
14: qb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_on:
20: em_or:
21: em_ov:
22: em_pa:
23: em_ph:
24: em_pl:
25: em_fo:
26: em_sts:
27: em_sy:
28: em_vl:
29: em_higo_hum:
30: em_higo_inv:
31: em_htgo_rod:
32: em_htgo_inv:
33: em_htgo_hum:
34: em_higo_inv:
35: em_htgo_rod:
36: em_higo_other:

Result No.	Score	Query	Match	Length	DB	ID	Description
c	1	1.012	72.9	31049	3	U55376	U55376 Caenorhabdi
c	2	90.8	6.5	13129	2	AC019610	AC019610 Drosophil
c	3	90.8	6.5	64674	3	AC004247	AC004247 Drosophil
c	4	90.8	6.5	174316	3	AC092230	AC092230 Drosophil
c	5	90.8	6.5	259164	3	AB03614	AB03614 Drosophil
c	6	84.8	6.1	31053	3	CIEP4D12	268298 Caenorhabdi
c	7	71.2	5.1	40664	3	CIE4TA4	282263 Caenorhabdi
c	8	70.8	5.1	25560	3	CEZB09	AC03170 Caenorhabdi
c	9	68	4.9	7218	6	I66494	Sequence 14
c	10	65.8	4.7	34017	3	CE09A9	279601 Caenorhabdi
c	11	47.8	3.4	253305	3	PFMAL3P7	AL034559 Plasmodiu
c	12	44.6	3.2	3885	3	AFL151733	AL034559 Plasmodiu
c	13	44.6	3.2	3985	3	AFL151111	AF151111 Dictyoste
c	14	44.6	3.2	110211	8	AC005170	AC005170 Arabidops
c	15	43.8	3.2	366	8	X59536	S.cerevisia
c	16	43.8	3.2	21468	8	M11449	Yeast mitoc
c	17	43.8	3.2	21154	8	YSCMTCG13	L36897 Saccharomy
c	18	43.8	3.2	85779	8	AJ011856	AJ011856 Saccharom
c	19	43.8	3.2	112576	2	AC011250	AC011250 Drosophil
c	20	43.8	3.2	184610	3	AC074440	AC074440 Drosophil
c	21	43.8	3.2	262231	3	AE003823	AE003823 Drosophil
c	22	43.4	3.1	77835	2	PFMAL13P2_3	Continuation (4 of
c	23	43	3.1	1744	6	AX083744	AX083744 Sequence
c	24	43	3.1	33637	3	CIE10G11	AL161711 Caeorhabd
c	25	42.8	3.1	17577	9	HS1033810	AL031228 Human DNA
c	26	42.4	3.1	21015	3	AFL19082	ALF19082 Haemato
c	27	42.4	3.1	321003	2	PFMAL4P3	AL035416 Plasmmodiu
c	28	42.2	3.0	104992	2	AC005504	AC005504 Plasmmodiu
c	29	42.2	3.0	169516	2	AC004157	AC004157 Plasmmodiu
c	30	42	3.0	23269	8	SPAC1856	AL391783 S.pombe c
c	31	42	3.0	23269	8	SPAC1856	AL391783 S.pombe c
c	32	41.6	3.0	175338	2	AC084352	AC084352 Homo sapi
c	33	41.6	3.0	321003	2	PFMAL4P3	AL035416 Homo sapi
c	34	41.2	3.0	64248	2	AC017951	AC017951 Drosophil
c	35	41.2	3.0	84055	3	AC004439	AC004439 Drosophil
c	36	41.2	3.0	170868	3	AC092221	AC092221 Drosophil
c	37	41.2	3.0	30652	3	AE003586	AE003586 Drosophil
c	38	41	3.0	3115	3	ECAG270225	AJ270225 Entodiniu
c	39	40.8	2.9	178386	2	AC073485	AC073485 Homo sapi
c	40	40.8	2.9	191122	9	AC012494	AC012494 Homo sapi
c	41	40.6	2.9	14933	3	AE001394	AE001394 Plasmmodiu
c	42	40.6	2.9	88512	8	F21H2	AC007894 Arabidops
c	43	40.6	2.9	199551	8	AC018460	AC018460 Arabidops
c	44	40.6	2.9	199551	2	AC006281	AC006281 Plasmmodiu
c	45	40.4	2.9	2206	8	SCYKR072C	228297 S.cerevisia

ALIGNMENTS

RESULT	1	LOCUS	U55376/C	DEFINITION	Caenorhabditis elegans cosmid F16H11, complete sequence.
REFERENCE	1	VERSION	U55376.1	ACCESSION	09-AUG-2001
AUTHORS		TITLE			
					The C. elegans Sequencing Consortium.
					Genome sequencing of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium. Science. 282 (5396), 2012-2018 (1998)
JOURNAL					
MEDLINE					
REFERENCE	2				
AUTHORS					
TITLE					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruityfly.org/sequence/>) or send email to bddp@fruitfly.berkeley.edu.
Location/Qualifiers

It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone F44D12 is at 1 in this sequence. The true right end of clone F44D12 is at 6268 in sequence Z68299.

The true left end of clone T04B2 is at 31750 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence 268299.

The end of this sequence (31750..31855) overlaps with the start of sequence 268299.

For a graphical representation of this sequence and its analysis see: <http://wwwbase.sanger.ac.uk/peri/ace/elegans/seq/sequence?name=F44D12>.

FEATURES

source

1. 31853 Location/Qualifiers

/organism="Caenorhabditis elegans"

/db_xref="taxon:6239"

/chromosome="IV"

/clone="F44D12"

join(3277..3413,3774..3951,3994..4239,4318..4603,

4647..4884,4949..5180,5225..5395,5449..5759,5805..6053,

6100..6802,6846..7020,7211..7306,7357..7425)

/gene="F44D12_1"

join(3277..3413,3774..3951,3994..4239,4318..4603,

4647..4884,4945..5180,5225..5395,5449..5759,5805..6053,

6100..6802,6846..7020,7211..7306,7357..7425)

/gene="F44D12_1"

/note="Similarity to Mouse Tight junction protein 20-1

(SW:201_MOUSE), contains similarity to Pfam domain:

PF00295 (PDZ domain (Also known as DHR or GLGF).)

Score=35.0, E-value=5.5e-07, N=2

cDNA EST Yk16d12.5 comes from this gene

cDNA EST Yk45f4.5 comes from this gene

cDNA EST Yk129e3.3 comes from this gene

cDNA EST Yk129e3.5 comes from this gene

cDNA EST Yk360b12.3 comes from this gene

cDNA EST Yk352d5.3 comes from this gene

cDNA EST Yk352d5.5 comes from this gene

cDNA EST Yk324a9.3 comes from this gene

cDNA EST Yk324a9.5 comes from this gene

cDNA EST Yk219a4.3 comes from this gene

cDNA EST Yk219a4.5 comes from this gene

cDNA EST Yk218f7.3 comes from this gene

cDNA EST Yk218f7.5 comes from this gene

cDNA EST Yk220g3.5 comes from this gene

cDNA EST Yk252a9.5 comes from this gene

cDNA EST Yk357f4.5 comes from this gene

cDNA EST Yk388f10.5 comes from this gene"

/codon_start=1

/protein_id="CAA92607.1"

/db_xref="GI:387158"

/db_xref="SPTRMBL:020398"

/translation="MEKEIIEVEAKWMEKKVKAERDDGKIDFEEKIDSLOIQV

HRSEYTKMEEEANKRKLRLQEALORDYEETMQERSIVLEENSNEERQLKETEK

MKEDSEVAALKOELNTTRILKVNMEETQANAREAAARLTIAEQNARARITED

EARLICQQTGDAIIVDIWNTHSVQNLPIFKPGLMIGLQPTDGSIVGPIYRQI

AHGSPEDNVKFLDHIMVNDISVTMDERSVMGMISNCIIHLVIRRSCNKLSDV

CLPLANGVGEELSNGEVFINECEPNGAASREGLAPEQRVWVHMTPYDAHQAEMIJKN

SREPLVIGTILOSTKGDHNGDKHQTIERFWESNGGSSEKERTVYAKANIDRSNDQV

LLRQCSLRMFAQSPAMSPLVYGSLRAPYSSLDHTKMLDALKRFAKSTAHSE

ISSSSTAAPALWPSPSVNNKDLVPAFTGATNEGVPPVKTIFSPVTRLIRONS

EOKRPMQSOISRHGENDGRCFSNSNPTSDRPRSMHETPSSITMEGKPRRSAYVSP

SHPPYPTIRDMSMSSMSSNSIRLPSTSFSINQPNNSCSSLTGSGVPPRHCJSD

GSDFSLSTTSAMHSYAEGRSNKYLHPRONHHEPSLLVNRSSQERLYVPRDSV

GGNAIGTIAKSIGSDLILLEDGLLSDIDGCVNVRNTLECAMNTSLADNSEJGLLWQDG

GDRLRLRLAGDGSFFLRNIIDRMENKOBELIKCGDVFVDTMLMGTRGRWAKW

VDKECROREGAIPSETTYQAIRANRYAMPFPEPKAYENWEKLOTKVKREVLLFGAVV

EPFLOMLVDESDKFSIVARRSLLDVKVLQDQSAIDLRLKRCRMFFILVKIREKSKYKOLKDQVNEUICGETTSSK

IMDITAQGLICLVCLOQDQSAIDLRLKRCRMFFILVKIREKSKYKOLKDQVNEUICGETTSSK

VDEE" complement(join(7556..7673,7717..7832,8096..8296,
8340..8510,8562..8718,8944..9111,9541..9638))
/gene="F44D12_2"
complement(join(7556..7673,7717..7832,8096..8296,
8340..8510,8562..8718,8944..9111,9541..9638))
/gene="F44D12_2"
/note="CDNA EST Yk16g12.3 comes from this gene
cDNA EST Yk167d4.3 comes from this gene
cDNA EST Yk53h4.5 comes from this gene
cDNA EST Yk53h9.5 comes from this gene
codon_start=1
/codon_id="CAA92598..2"
/protein_id="GI:5824507"
/db_xref="SPTRMBL:Q20392"
/translation="MOTLILQILVLFTHFVNLYTCSCRGNEBNFNSRSTCRGD
FCYHAQFQYGNMMPYIQKQGYTGEVAPEGCDEVHNHVNMCPCSDADYCNAINTSLLAD
TATTLLPVOTCPEKVNMPKPRNTKCAANVCTEFLAKTOPEVNSNANTWTWKDCDNK
ENBFDFEFPTLTFNPFEGTCVWLNQGSDQPDTHACYGSLSLQPTLVFLDFVTTACTHVGD
YENPRLQYVKSOSSCGOFCFSATSRSGEVFGCVNSVTEVGATQPLKTYTRAYTGV
QWICDQSYCNADLKSAELSWSPPELYLXRNISNREFNVFYIDSARSSSSIFIPLAIPIVF
LINYCFH"
join(12326..12380,12427..12481,12574..12895)
/gene="F44D12_9"
join(1..12380,12427..12481,12574..12895)
/gene="F44D12_9"
/codon_start=1
/protein_id="CAA92605..1"
/db_xref="GI:3877156"
/db_xref="SPTRMBL:Q20396"
/translation="MTHMHQEEQEGSPLSISYFOYLHVEYNICGRKGWGTOTTLMKNEELA
ENSSVTONSQIGLVTPTDWTGKREKSPSPIDRWLVYIIFTMHONGMMSWNMFITAPQ
YYHDYWINNTNYQDSEWSNLTGTSQENVGTMILNTTIVMV"
complement(115660..15889)
/gene="F44D12..3"
complement(115660..15889)
/note="Contains similarity to Pfam domain: PF00635 (MSP
(Major sperm protein) domain), Score=111.7,
E-value=4.7e-30, N=1"
cdNA EST Yk57e8.3 comes from this gene; cDNA EST yk102a1..3
comes from this gene
cdNA EST Yk57e8.5 comes from this gene; cDNA EST yk170h1..5
comes from this gene
cdNA EST Yk122a5.5 comes from this gene; cDNA EST yk122a5
comes from this gene
cdNA EST Yk149d3..5 comes from this gene
cdNA EST Yk151a2..5 comes from this gene
cdNA EST Yk151a2..5 comes from this gene
cdNA EST Yk164g1..5 comes from this gene
cdNA EST Yk170h1..5 comes from this gene
cdNA EST Yk179d6..5 comes from this gene
cdNA EST Yk189f1..5 comes from this gene; cDNA EST
cdNA EST Yk151a2..3 comes from this gene; cDNA EST
cdNA EST Yk149d3..5 comes from this gene
cdNA EST Yk151a2..5 comes from this gene
cdNA EST Yk164g1..3 comes from this gene
cdNA EST Yk170h1..3 comes from this gene
cdNA EST Yk179d6..3 comes from this gene
cdNA EST Yk189f1..3 comes from this gene
cdNA EST Yk151a2..3 comes from this gene; cDNA EST
cdNA EST Yk149d3..5 comes from this gene
cdNA EST Yk2f10..3 comes from this gene; cDNA EST
Yk150a1..5 comes from this gene
cdNA EST Yk164f1..5 comes from this gene; cDNA EST
Yk378f3..3 comes from this gene
cdNA EST Yk378f3..3 comes from this gene; cDNA EST
Yk149d3..3 comes from this gene
Yk149d3..5 comes from this gene
Yk164g1..3 comes from this gene
Yk170h1..3 comes from this gene
Yk179d6..3 comes from this gene
cdNA EST Yk189f1..3 comes from this gene; cDNA EST
cdNA EST Yk431f3..3 comes from this gene; cDNA EST
cdNA EST Yk431f3..5 comes from this gene; cDNA EST
Yk365g9..3 comes from this gene
cdNA EST Yk365g9..5 comes from this gene
Yk311h6..3 comes from this gene
cdNA EST Yk311h6..5 comes from this gene
cdNA EST Yk311h6..5 comes from this gene; cDNA EST
Yk29b2..3 comes from this gene
cdNA EST Yk399b2..5 comes from this gene; cDNA EST
Yk309f7..3 comes from this gene

FEATURES	source
COMMENT	Washington University, St. Louis, MO 63110, USA. E-mail: jes@anu.ac.uk or rwenematode.wustl.edu On Nov 13, 1998 this sequence version replaced gi:2528885. Coding sequences below are predicted from computer analysis, using predictions from GeneFinder (P. Green, U. Washington), and other available information.
Current sequence finishing criteria for the <i>C. elegans</i> genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. (981006 d1) ; Cosmid flipped	
IMPORTANT: this sequence is not the entire insert of clone C47A4 . It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.	
The true right end of clone C47A4 is at 40664 in this sequence. The true left end of clone F52D4 is at 101 in this sequence. The true right end of clone F56F12 is at 9819 in this sequence. The start of this sequence (1 .. 104) overlaps with the end of sequence 280218. The end of this sequence (40564 .. 40664) overlaps with the start of sequence 282273.	
For a graphical representation of this sequence and its analysis see: http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=C47A4 .	
Location/Qualifiers	
1..40664	/organism="Caenorhabditis elegans" /db_xref="taxon:6239" /chromosome="IV" /clone="C47A4"
	complement(join(3491..3973,4192..4254))
	/gene="C47A4.1"
	complement(join(3491..3973,4192..4254))
	/gene="C47A4.1"
	/codon_start=1
	/protein_id="CAB62792.1"
	/db_xref="GI:6562350"
	/db_xref="SPTREMBL:Q9U3L9"
	/translation="MTIKDLKLIENFIIVTVPSETLEMDRNGKLKKNKGVNDDEV QIILDNLDVIGGYARESITDTLANTIIIFPTTIVCFWTAIWWYRRAEIQKEYDD KLLIRKLGCVSQEDDFQYTDIDDDLEPEBCWLTAKAERAAEQKMAQ RYKAVKRYKRNAGTISFVSDD"
gene	join(15731..15769,15884..16078,16580..16814,17751..18821..18923,18970..19055,19306..19502,20056..20176)
CDS	/gene="C47A4.2"
	join(15731..15769,15884..16078,16580..16814,17751..18821..18923,18970..19055,19306..19502,20056..20176)
	/gene="C47A4.2"
	/note="contains similarity to Pfam domain: PF01733 (Nucleoside transporter), Score=142.5, E-value=2.4e-39, N=1"
	/codon_start=1
	/protein_id="CAB62793.1"
	/db_xref="GI:6562351"
	/db_xref="SPTREMBL:Q9U3L9"
	/translation="MVTIDVEISIPVDPPPLYTDONTEDLASIVDGDSSDEHNE PEDGRIVEWILLINGVGVLPWNNTIAQQYYDQWFTWNGTATHYADSPQSAMM VAQVPNLIVAIINVNLIREPLPVRLVAPIAFNSLIVIILALVTFQPSDQVLAIVATI STQETEVPAFVLSLQLLQNSFEGMWMQCNMFLVCLVSNWFCRMWLOCICVFLVYFVLSVPTVLAV QPGYTVFPMDVYAGTAVENNEFEEAVGNAATFTEPGPRILIVCPVIRLFLFIPF FSNLPLHSPMGVLFTNNLFFEGNTLFLATSGYFSSLGMVTPRCPPESKLAGG SAULSVLQTAGTATMVNNT"
gene	complement(join(28627..28707,28755..28832,28884..29024, 29070..29267,29338..29466,29527..29650))
CDS	/gene="C47A4.3"
	complement(join(28627..28707,28755..28832,28884..29024, 29070..29267,29338..29466,29527..29650))
	/gene="C47A4.3"
	/note="contains similarity to Pfam domain: PF00149 (Ser/thr protein phosphatase), Score=476.0, E-value=1.39, N=1"

ORGANISM		COMMENT	
Unknown.		jesesanger.ac.uk or rwennemate.wustl.edu	
Unclassified.		On Sep 24, 1998 this sequence version replaced gi:1515146.	
REFERENCE	1 (bases 1 to 7218)	Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.	
AUTHORS	Dorner F., Scheifflinger, F. and Falkner, F. Gunter.	Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction. Subclones. Exceptions are indicated by an explicit note.	
JOURNAL	Recombinant fowlpox virus	IMPORTANT: This sequence is not the entire insert of clone K09A9. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.	
FEATURES		The true left end of clone K09A9 is at 35278 in sequence 279536.	
Best Local Similarity	4.2%	The true right end of clone K09A9 is at 34017 in this sequence. The true left end of clone K09B9 is at 24804 in this sequence. The true right end of clone C02C6 is at 104 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence Z79596. The end of this sequence (33914..34017) overlaps with the start of sequence Z79602.	
Matches	17;	For a graphical representation of this sequence and its analysis see: http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=K09A9 .	
Conservative	234;	Location/Qualifiers	
Patent	US 5670367-A 14 23-SEP-1997;	1..7218	
Source	/organism="unknown"	/organism="Caenorhabditis elegans"	
ASE COUNT	1 1944 a	/db_xref="taxon:6239"	
ORIGIN	1491 c	/clone="K09A9"	
	1486 g	/chromosome="X"	
	1929 t	/note="predicted using GeneFinder	
	368 others	Similarity to Bovine NADH-ubiquinone oxidoreductase 49 KD subunit (SW:NUC_BOVIN), contains similarity to Fam domain: PR00346 (Respiratory-chain NADH dehydrogenase, 49 Kd subunit), Score=55.8, E-value=6.4e-215, N=1	
FEATURES		CDNA EST BMBL:Z4772 comes from this gene	
source		CDNA EST EMBL:M88875 comes from this gene	
		CDNA EST EMBL:M89294 comes from this gene	
		CDNA EST EMBL:Y01723 comes from this gene	
		CDNA EST YK81b6_3 comes from this gene	
		CDNA EST YK81b6_5 comes from this gene	
		CDNA EST YK68h5_3 comes from this gene	
		CDNA EST YK68h5_5 comes from this gene	
		CDNA EST YK70h1_5 comes from this gene	
		CDNA EST YK70h1_3 comes from this gene	
		CDNA EST YK454f3_3 comes from this gene	
		CDNA EST YK454f3_5 comes from this gene	
		CDNA EST YK4859_3 comes from this gene	
		CDNA EST YK4859_5 comes from this gene	
		CDNA EST YK454f8_3 comes from this gene	
		CDNA EST YK454f8_5 comes from this gene	
		CDNA EST YK424c6_3 comes from this gene	
		CDNA EST YK424c6_5 comes from this gene	
		CDNA EST YK1845_3 comes from this gene	
		CDNA EST YK418a5_5 comes from this gene	
		CDNA EST YK229c10_3 comes from this gene	
		CDNA EST YK652a6_3 comes from this gene	
		CDNA EST YK68098_3 comes from this gene	
		CDNA EST YK530c9_5 comes from this gene	
		CDNA EST YK530c9_3 comes from this gene	
		CDNA EST YK532b10_3 comes from this gene	
		CDNA EST YK596h3_3 comes from this gene	
		CDNA EST YK652a6_3 comes from this gene	
		CDNA EST YK68098_3 comes from this gene	
		CDNA EST YK530c9_5 comes from this gene	
		CDNA EST YK562f5_3 comes from this gene	
		/codon_start=[1]	
REFERENCE	CEK09A9	34017 bp DNA	
AUTHORS	Caenorhabditis elegans cosmid K09A9,	INV 20-JUN-2001	
JOURNAL	complete sequence.	Genome sequence of the nematode <i>C. elegans</i> ; a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium Science 282 (5396), 2012-2018 (1998)	
FEATURES		The <i>C. elegans</i> sequencing consortium.	
DEFINITION		The <i>C. elegans</i> sequencing consortium.	
ACCESSION		The <i>C. elegans</i> sequencing consortium.	
VERSIONS		The <i>C. elegans</i> sequencing consortium.	
KEYWORDS		The <i>C. elegans</i> sequencing consortium.	
SOURCE		The <i>C. elegans</i> sequencing consortium.	
ORGANISM		The <i>C. elegans</i> sequencing consortium.	
MEDLINE		The <i>C. elegans</i> sequencing consortium.	
REMARK		The <i>C. elegans</i> sequencing consortium.	
REFERENCE	1 (bases 1 to 34017)	none.	
AUTHORS		Genome sequence of the nematode <i>C. elegans</i> ; a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium Science 282 (5396), 2012-2018 (1998)	
JOURNAL		Submitted (29-AUG-1996) Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10 1RQ, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:	
TITLE		Direct Submission	
		Submitted (29-AUG-1996) Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10 1RQ, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:	

CLLSVENNYYKKQKVKSYSLFLGSLPSEGPGCKGLKKNAPLYNKBNKQYGNPFHLL
YYPPVTEESDKNEERNEERNEERNEERNEERNEERNEERNEERNEERNEERNEER
NGTQSYHHNNNSINSNDDOKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
YENDININKQFNFNIVDLSFNITNEWYSEFVLNFNIVDLSFNITNEWYSEFVLNFN
KNLPLPTDIEPKMNSKINAEDTENRFLYVVECYLCEVNCERKKGKTYELEK
LNDFFETKQDMEKTMNSKINAEDTENRFLYVVECYLCEVNCERKKGKTYELEK
CCLKINKYYYSWLLFYELFLFENITFLTVYVYKLMNFILFLFLYFLMNS
DLFTFLCMPSMNSNSINYATFLFLCMPSMNSLTSIHSASVNLTEFLIPPSKFCMS
MLNTRKGRSKSLTDKRKSSQTYEGNINMDMSXYPTHOLKDNSDFMLEGSDBD
DNTTNTNN
NGSCHSTLNITRMSKOSGEINFKNYSKRNKDTMDVNKNRNLGRCSNSY
IKEFMRVTRKRNNDNTYDNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
DLDFLTSYSONNLYENTUYETIKLFLYLYNNKNDYLTNKRTKLTMNDLVOY
NDQKLNMDMEDEKKVKSIFCFLYKRDYLDEPTEAIDKTKTFLKFDRKKNNT
FICHTYRANNFLANDKVGSKSFKQLFQKSYKETFLNQGPNPEOECND
YQSERKUMNLLKQDQGSSSYWTFRFLPFLSLIKVSLGQYQNTD
NKNKNSVHDIDHPLITVNTNEFTNNRDIKRNIKGTRNENRNCIFIPNENN
YCTKINIELESLKCLCUNKFKNLTQLKTDIYTUYITYTLNKKKKLKKNIODK
DKYLETDLPLFLFQNYFLNELNLMKNNHNNNNNNNNNNNNNNNNNNNN
LINSTNHDDPYDINMHNHNNHKYKQYKUNNNNNNNNNNNNNNNNNNN
KETKEFSTSYNRNKNNDENDEGYEYDNMMSNTSNNMKKKKKKKKKKK
KETKEFSTSYNRNKNNDENDEGYEYDNMMSNTSNNMKKKKKKKKKKK
NPKFETTYKPLLKKLGQSSSYWTFRFLPFLSLIKVSLGQYQNTD
YSTLSSNNHLQSTINYFLYKNNNNNNNLTGYNQNLDDNKTNNLNNLNN
TKODDNNAEISHINNNDHNDNIDKDLKLENNEYANXKLJHENESNYYAISTINSLORYC
LKEYTINYIDEPLEDDNLSINNLYLEKAKKNTIISLGFPAFLITERDNNKED
ETLETRDNPKEVNEINLYNLYTSHSYAYNTSNSLNNMSDFQNLNNKSGNT
ILDGSTSYDNKUVVEDYKNCMDMTLYDENKNYKLIDDNKNTRECNAFLRSK
VSYNNNEEEEDINGNINDNNNNITPKNNKNEEEDINGNINDNNNN
FHEFMDRQDINYFLSIVIIVFLCFFCVERKNTENRKFENFHVOHYQOILLE
YLYFFFLYUCLCFLTIVYIDYKEFLMSFCELLYGENFIELSCLCLESYLYHSYLYFL
FLFLFICGFLISIVIIVLSSYANLFLNLSHVFLFLFDSFAHVLNIRSC
LNUVRHMKLDEDIMDMSNNNNFVIGCFCKVYQDLSGDNLNTQGSCVDES
SFFNTSGDVFLINCIYLISIVYKLQJLPSKERTKKNKSEKTYEEIINNSY
VGWORKKCNKFRKNTFLYKFLNKSKEPKRNPFFYIQLFLNUKGRNN
DNTTNTNNMGGYKGVKSDNQSEKLNKEDYNNLSEKTNQNLNQSEKLNKEDYNN
NTVDEMOQSEKSNIKEMLYKETEIDYKKGSDDDEENKNKLSQKYLTKLNKLN
KMRPQKTPSSLFNDDNLYFLYFICFNKONLGKQDLAGYQONEEINNKFK
NONBHQESTKQKDNKTRKJELPTMSYEHPKLTFLKTENLVEYKLTINMLK
VNLAQDIIHNCNOLSGMMKKKKVFLMINLRDKLFLYKLNKNDIDFCQSYQYINILKLN

Query	Match	3.4%	Score	47.8;	DB	3;	Length	253305;
Best Local	Similarity	61.88;	Pred.	No.	0.25;			
Matches	Mismatches	0;	Indels	47;	Gaps			
1266	tgttattggactctcgggacaaggccaaqtatcttataaatatttagatqatagatqat	132						
78852	TTTATCTTCATTGTTAAAAGATGTATTTAATATTATAATTGTATAATA	78852						
1326	tacttgttatgttttatataaactgtgtgaaataaaataattataaaaaaa	1386						
78912	TAATTTATCTTCATTGTTAAAAGATGTATTTAATATTATAATTGTATAATA	78912						

	RESULT	12	Qua lity	Bee Ma	Ma	RESU LT
1386	aaa	1388				
78972	ATA	78974				
DEFINITION	AF151733	3885 bp	DNA	INV	07-MAY-2001	Db
CUS	AF151733	Dicyostelium discoideum	F-Box A protein (fbxA)	gene, complete cds.		Qy
VERSION	AF151733.1	GI:5163417				Db
KEYWORDS	Dictyostelium discoideum					
SOURCE	Dictyostelium discoideum					
ORGANISM	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium					
REFERENCE	1 (bases 1 to 3885)					
AUTHORS	Nelson M K Clark A					

and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Liy Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

Location/Qualifiers

1..110211

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="II"

/note="Sequence from clone F27L4"

repeat_region

complement(<1..2793)

/rpt_family="TAAAn"

complement(356..401)

/rpt_start=1

/rpt_end=3012)

/note="Sequence from clone F27L4"

gene

complement(join(1549..2058, 2731..3012))

/rpt_start=1

/rpt_end=3012)

/note="Sequence from clone F27L4"

gene

complement(join(1744..2058, 2731..3003))

/rpt_start=1

/rpt_end=3003)

/note="Sequence from clone F29E15"

repeat_region

complement(3567..3677)

/rpt_family="TAAAn"

complement(3570..3641)

/rpt_start=1

/rpt_end=3641)

/note="Sequence from clone F29E15"

gene

complement(4565..4612)

/rpt_start=1

/rpt_end=4612)

/note="Sequence from clone F29E15"

mRNA

join(<6480..6749, 7602..7726, 7811..7848, 7929..8039,

repeat_region

complement(6480..6749)

/rpt_start=1

/rpt_end=6749)

/note="Sequence from clone F29E15"

gene

complement(6480..8268)

/rpt_start=1

/rpt_end=8268)

/note="Sequence from clone F29E15"

gene

complement(6480..8786)

/rpt_start=1

/rpt_end=8786)

/note="Sequence from clone F29E15"

gene

complement(6480..8786)

/rpt_start=1

RESULT		15				
	SCORI7/c		SCORI7	366 bp	DNA	PLN
DEFINITION	S.cerevisiae DNA for ori 7.					26-FEB-1992
ACCESSION	X59536					
VERSION	X59536.1					
KEYWORD	GI:14 234					
SOURCE	Baker's yeast.					
ORGANISM	Mitochondrion <i>Saccharomyces cerevisiae</i>					
Bdukaryote; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces .						
REFERENCE	1 (bases 1 to 366)					
AUTHORS	Skelly P.J. and Clark-Walker G.D.					
TITLE	Sequence rearrangements at the ori 7 region of <i>Saccharomyces cerevisiae</i> mitochondrial DNA					
JOURNAL	J. Mol. Evol.					
MEDLINE	32 (5), 439-442 (1991)					
FEATURES	91251143					
source	Location/Qualifiers					
	1. .366					
	/organism="Saccharomyces cerevisiae"					
	/organelle="mitochondrion"					
	/db_xref="taxon:4932"					
BASE COUNT	188	a	25	c	28	g
BASE CNT					125	t

Search completed: February 28, 2002, 00:56:47
Job time: 5549 sec

THIS PAGE BLANK

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: February 27, 2002, 23:59:59 ; Search time 113.25 Seconds

(without alignments)
10507.439 Million cell updates/sec

Title: US-08-816-011f-36
Perfect score: 1388

Sequence: 1 atggtaataatccaggatc.....ttattaaaaaaa 1388

scoring table: IDENTITY_NUC GapP 10.0 , Gapext 1.0

Searched: 930521 seqs, 428662619 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing First 45 summaries

Database : N_Geneseq_1101;*

1: /SIDS2/gcdata/geneseq/geneseq/NA1980.DAT:*

2: /SIDS2/gcdata/geneseq/geneseq/NA1981.DAT:*

3: /SIDS2/gcdata/geneseq/geneseq/NA1982.DAT:*

4: /SIDS2/gcdata/geneseq/geneseq/NA1983.DAT:*

5: /SIDS2/gcdata/geneseq/geneseq/NA1984.DAT:*

6: /SIDS2/gcdata/geneseq/geneseq/NA1985.DAT:*

7: /SIDS2/gcdata/geneseq/geneseq/NA1986.DAT:*

8: /SIDS2/gcdata/geneseq/geneseq/NA1987.DAT:*

9: /SIDS2/gcdata/geneseq/geneseq/NA1988.DAT:*

10: /SIDS2/gcdata/geneseq/geneseq/NA1989.DAT:*

11: /SIDS2/gcdata/geneseq/geneseq/NA1990.DAT:*

12: /SIDS2/gcdata/geneseq/geneseq/NA1991.DAT:*

13: /SIDS2/gcdata/geneseq/geneseq/NA1992.DAT:*

14: /SIDS2/gcdata/geneseq/geneseq/NA1993.DAT:*

15: /SIDS2/gcdata/geneseq/geneseq/NA1994.DAT:*

16: /SIDS2/gcdata/geneseq/geneseq/NA1995.DAT:*

17: /SIDS2/gcdata/geneseq/geneseq/NA1996.DAT:*

18: /SIDS2/gcdata/geneseq/geneseq/NA1997.DAT:*

19: /SIDS2/gcdata/geneseq/geneseq/NA1998.DAT:*

20: /SIDS2/gcdata/geneseq/geneseq/NA1999.DAT:*

21: /SIDS2/gcdata/geneseq/geneseq/NA2000.DAT:*

22: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1386.4	99.9	1388	17 AAT18168	CORK potassium channel gene
2	85	6.1	936	22 AAF58252	Oligonucleotide D1
3	85	6.1	936	22 AAF58254	Oligonucleotide D1
4	85	6.1	936	22 AAF58257	Oligonucleotide D1
5	85	6.1	936	22 AAF58259	Oligonucleotide D2
6	85	6.1	936	22 AAF58262	Oligonucleotide D2
7	85	6.1	938	22 AAF58255	Oligonucleotide D1
c 8	83	6.0	936	22 AAF58252	Oligonucleotide D1
c 9	83	6.0	936	22 AAF58254	Oligonucleotide D1
c 10	83	6.0	936	22 AAF58257	Oligonucleotide D1
c 11	83	6.0	936	22 AAF58259	Oligonucleotide D2

ALIGNMENTS

27	38	2.7	4709	18 AAC7990
28	37	2.7	2079	AAA26360
c 29	36.8	2.7	67792	AAFB854
c 30	36.6	2.6	513445	AAI6133
c 31	36.4	2.6	1338	AAAC9411
c 32	36.2	2.6	1667	AAFB85601
c 33	36.2	2.6	1851	AAH33171
c 34	36.2	2.6	3076	AAH46551
c 35	36.2	2.6	5761	AAI59388
c 36	36	2.6	1077	AAV39005
c 37	36	2.6	21	AA37772
c 38	36	2.6	1164	AAV07654

c 39	36	2.6	1179	22 AAH33171
c 40	36	2.6	2215	AAQ70882
c 41	36	2.6	2621	AAQ76109
c 42	36	2.6	2916	AAQ11555
c 43	36	2.6	2916	AAQ89646
c 44	36	2.6	2964	AAFB8022

c 45	36	2.6	2970	20 AAX8641
------	----	-----	------	------------

RESULT 1	AAT18168	standard	cdDNA	1388 BP.
ID	AAT18168			
XX				
AC	AAT18168;			
XX				

DT	16-OCT-1996	(first entry)		
XX				
DE				
XX				
KW				
XX				
OS				
XX				

PH	Key CDS			
FT	/*tag= a			
FT	/product= CORK potassium channel gene			
FT	/*tag= b			
XX	W09613520-A1.			
XX				
PD	09-MAY-1996.			
XX				
PF	25-OCT-1995;			
XX	95WO-US14364.			
PR	31-OCT-1994;			
XX	94US-0332312.			
PA	(AMCY) AMERICAN CYANAMID CO.			
XX				

PN	W09613520-A1.			
XX				
PD	09-MAY-1996.			
XX				
PF	25-OCT-1995;			
XX	95WO-US14364.			
PR	31-OCT-1994;			
XX	94US-0332312.			
PA	(AMCY) AMERICAN CYANAMID CO.			
XX				

XX 26-JUL-2000; 2000WO-US20476.
 XX PR 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190359.
 XX PA (CLIN) CLINICAL MICRO SENSORS INC.
 XX PI Umek RM;
 XX DR WPI; 2001-159728/16.
 XX Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 XX PS Example 6; Page 127; 159pp; English.
 XX CC The present invention relates to a composition comprising two nucleic
 acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 detection of nucleic acids, especially of substitutions (mismatches)
 CC and single nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 6.1%; Score 85; DB 22; Length 936;
 Best Local Similarity 0.9%; Pred. No. 8e-13; Mismatches 301; Indels 0; Gaps 0;
 Matches 7; Conservative 431; Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 XX DR WPI; 2001-159728/16.
 XX CC The present invention relates to a composition comprising two nucleic
 acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 detection of nucleic acids, especially of substitutions (mismatches)
 CC and single nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Example 6; Page 127; 159pp; English.
 XX DR WPI; 2001-159728/16.
 XX CC The present invention relates to a composition comprising two nucleic
 acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 detection of nucleic acids, especially of substitutions (mismatches)
 CC and single nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 6.1%; Score 85; DB 22; Length 936;
 Best Local Similarity 0.9%; Pred. No. 8e-13; Mismatches 301; Indels 0; Gaps 0;
 Matches 7; Conservative 431; Nucleic acids containing electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 detection of nucleic acids, especially of substitutions (mismatches)
 CC and single nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.

Search completed: February 28, 2002, 00:52:31
Job time: 3152 sec

Result No.	Score	Query Match	Length	DB ID	Description	SUMMARIES	
c 1	86.6	6.2	533	10 AU222343	AU222343 mRNA	FEATURES	8
c 2	77.6	5.6	332	10 AU201199	AU201199 mRNA	Source	
c 3	75.6	5.4	710	10 AU216992	AU216992 mRNA		
c 4	75.2	5.4	648	10 AW453425	AW453425 mRNA		
c 5	73.8	5.3	327	10 AU109268	AU109268 mRNA		
c 6	73.8	5.3	642	10 AI082929	AI082929 mRNA		
c 7	70.8	5.1	495	10 AU209436	AU209436 mRNA		
c 8	70.2	5.1	360	11 C13829	C13829 mRNA		
c 9	67.8	4.9	300	10 AU113471	AU113471 mRNA		
c 10	57.6	4.1	300	10 AU115402	AU115402 mRNA		
c 11	57.4	4.1	300	10 AU114995	AU114995 mRNA		
c 12	56.2	4.0	997	13 CNS005TE	CNS005TE mRNA		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	86.6	6.2	533	10 AU222343	AU222343 mRNA
c 2	77.6	5.6	332	10 AU201199	AU201199 mRNA
c 3	75.6	5.4	710	10 AU216992	AU216992 mRNA
c 4	75.2	5.4	648	10 AW453425	AW453425 mRNA
c 5	73.8	5.3	327	10 AU109268	AU109268 mRNA
c 6	73.8	5.3	642	10 AI082929	AI082929 mRNA
c 7	70.8	5.1	495	10 AU209436	AU209436 mRNA
c 8	70.2	5.1	360	11 C13829	C13829 mRNA
c 9	67.8	4.9	300	10 AU113471	AU113471 mRNA
c 10	57.6	4.1	300	10 AU115402	AU115402 mRNA
c 11	57.4	4.1	300	10 AU114995	AU114995 mRNA
c 12	56.2	4.0	997	13 CNS005TE	CNS005TE mRNA

AA406898 MBARC27F00
AU09482 SWOY3MCAM
AU109592 ALA30789 clone XBAI
AL054622 Drosophil
AA842090 MBACF3B0
AI784877 SWAMC4A0
AI239052 GH15222.5
R03450 PK06903.R1
CA3216 CA3216 Yuji
BB181657 BB181657
AU111089 AU111089
AU111161 AU111161
AQ842781 CGI1299A
AQ254298 CG0751B
C44886 C44886 Yuji
AU111560 AU111560
AU111734 AU111734
AQ254261 CG0732A
AW839621 ra4id09.Y
AW839541 ra4id09.Y
AU131074 mj58509.X
BF459052 UI-M-BZL-
AIS95433 mj58509.Y
AA058117 mj58509.X
AW333515 S22G3 AGS
AW048063 UI-M-BH1-
ALI01356 Drosophil
AK050136 Mus muscu
BF460904 UI-M-CG0P
BB169820 BB169820
BB182391 BB182391

ALIGNMENTS

Scoring table: IDENTITY_NUC
Gapov 10.0 , Gapext 1.0
Title: US-08-816-01F-36
Perfect.score: 1388
Sequence: 1 atggtaataatccacgatc.....ttattaaaaaaaaaaa 1388
Run on: February 27, 2002, 23:21:59 ; Search time 1212.08 Seconds
(without alignments)
12305.409 Million cell updates/sec

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Database : Maximum Match 100%
Listings first 45 summaries

EST: *
1: em_est:fun:*

2: em_esthun:*

3: em_estin:*

4: em_estom:*

5: em_estpl:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_htc:*

10: gb_est1:*

11: gb_est2:*

12: gb_htc:*

13: gb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pln:*

18: em_gss_pro:*

19: em_gss_rod:*

20: em_gss_vrt:*

21: em_gss_other:*

RESULT 1
AU222343/C
LOCUS AU222343 unpublished oligo-capped cDNA library, stage EST
DEFINITION Caenorhabditis elegans cDNA clone YK1015b04 3', mRNA sequence.
ACCESSION AU222343
VERSION AU222343.1
KEYWORDS EST,
SOURCE Caenorhabditis elegans
ORGANISM Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditis
; Rhabditidae; paleodirinae; Caenorhabditis
REFERENCE 1 (bases 1 to 533)
AUTHORS Kohara, Y., Shin-i, T., Thierry-Mieg, J., Suzuki, Y., and Sugano, S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES 1. 533
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1015b04"
/clone_1lb="unpublished L1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

AUTHORS	Kohara, Y., Shin-i, T., Thierry-Mieg J., Thierry-Mieg, D., Suzuki, Y and Sugano, S.
TITLE	A complementary view of the <i>C. elegans</i> genome
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp.
FEATURES	Location/Qualifiers
	1. .327 /organism="Caenorhabditis elegans" /strain="N2" /db_xref="taxon:6239" /clone="yK708A4" /clone_1lb="unpublished oligo-capped cDNA library" /sex="Hermaphrodite" /tissue_type="whole animal" /dev_stage="varied"
BASE COUNT	93 a 66 c 72 g 96 t
ORIGIN	
RESULT	6
AI082929	AI082929 642 bp mRNA EST 13-AUG-1998 LOCUS SWAMC12G09SK Brugia malayi adult male CDNA (SAW94NL-BmAM) Brugia malayi cDNA clone SWAMC12G09 5', mRNA sequence.
DEFINITION	
ACCESSION	AI082929
VERSION	AI082929..1 GI:3419555
KEYWORDS	EST.
ORGANISM	Brugia malayi Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.
REFERENCE	1 (bases 1 to 642) Williams, S.A.
AUTHORS	Genes expressed in adult males of <i>Brugia malayi</i> .
TITLE	Unpublished (1995)
JOURNAL	Contact: Steven A. Williams
COMMENT	Molecular Parasitology Smith College Department of Biological Sciences Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA Tel: 4133853866 Fax: 4133853766 Email: genome@smith.edu Seq Primer: PBluescript SK.
FEATURES	Location/Qualifiers
	1. .642 /organism="Brugia malayi" /strain="IRS Labs" /db_xref="taxon:62279" /clone="SWAMC12G09"
source	

and DNA pol I. The library was constructed in the lambda *X*-zap XR vector and has 1×10^6 independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Litzow-Waniekwi in the laboratory of Dr. S. A. William Lustigman (email: sara.lustigman@yale.edu). The library is available from Dr. Sara Lustigman (email:

BASE COUNT 192 slustigm@nybc.org."
ORIGIN a 119 c 120 g 225 t

FEATURES
source Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp
 Location/Qualifiers
 .1. 345
 /organism="Caenorhabditis elegans"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="K7094"
 /clone_lib="unpublished oligo-capped cDNA library"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="varied"
BASE COUNT 100 a 86 c 91 g 1 t 1 others
ORIGIN

	Query Match	Best Local Similarity	Score	DB 10:	Length
Matches	69; Conservative	69.0%	3.6%; Score 50.4;	Pred No. 2;	345;
				Mismatches	0;
				Indels	0;
				Gaps	0;
Oy	52	c caa gag aca gta caat atgt tact gtg ctgc attc ttgt gatt ccgg agtt ttctt	111		
Db	106	CCAGAGACAGGGAAATCTGTGTTCTCATTTAATCTGCCTCATGGAATCGGAACTCTT	165		
Oy	112	ctgccatggaaatgttcatttatcgccccgttggat	151		
Db	166	ATGCCATGGAAATGCTTATACGATCTCTATGATT	205		

Search completed: February 28, 2002, 00:25:49
Job time: 3830 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: February 27, 2002, 23:24:08 ; Search time 49.03 Seconds
Title: US-08-816-011F-36
Perfect score: 1388
Sequence: 1 atgtaatacAACGAT.....ttataaaaaaaaaa 1388

Total number of hits satisfying chosen parameters: 702406
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post processing: Minimum Match 10%
Maximum Match 10%
Existing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/2/ina/PCUTS_COMB.seq:
6: /cgn2_6/ptodata/2/ina/backfiles.1.seq:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

702406

8

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	68	4.9	7218	1 US-08-232-163-14	Sequence 14, Appli	
2	38.6	2.8	1582	3 US-08-245-196B-10	Sequence 10, Appli	
3	38.6	2.8	1582	3 US-08-345-196B-12	Sequence 12, Appli	
4	36	2.6	1164	2 US-08-794-796-1	Sequence 1, Appli	
5	36	2.6	2621	2 US-08-553-619B-8	Sequence 8, Appli	
c	6	2.6	5852	1 US-07-867-106-2	Sequence 2, Appli	
7	35.8	2.6	1400	1 US-08-164-164-1	Sequence 1, Appli	
8	35.8	2.6	1400	1 US-08-338-057-1	Sequence 1, Appli	
9	35.8	2.6	1400	2 US-08-166-816-1	Sequence 1, Appli	
10	35.4	2.6	581	2 US-08-557-309B-22	Sequence 22, Appli	
11	35.4	2.6	581	3 US-08-834-306-22	Sequence 22, Appli	
12	35.4	2.6	581	4 US-08-593-674A-22	Sequence 11, Appli	
c	13	35.4	2.6	1303	2 US-08-793-410-11	Sequence 1, Appli
c	14	35.4	2.6	3489	2 US-08-72-323A-1	Sequence 1, Appli
c	15	35.4	2.6	32207	4 US-08-757-669A-20	Sequence 20, Appli
c	16	35.4	2.6	32207	4 US-08-757-669A-20	Sequence 20, Appli
c	17	35.2	2.5	972	3 US-09-286-650-1	Sequence 1, Appli
c	18	2.5	51259	3 US-08-181-891-209	Sequence 20, Appli	
c	19	34.8	2.5	289	4 US-09-007-005-17	Sequence 17, Appli
c	20	34.8	2.5	289	4 US-09-244-796-17	Sequence 17, Appli
c	21	34.6	2.5	991	4 US-09-377-648-7	Sequence 1, Appli
c	22	34.6	2.5	2238	1 US-08-742-011-1	Sequence 1, Appli
c	23	34.6	2.5	3645	2 US-08-663-112-1	Sequence 1, Appli
c	24	34.2	2.5	852	4 US-09-161-691-1	Sequence 1, Appli
c	25	34.2	2.5	1991	2 US-08-415-533-40	Sequence 1, Appli
c	26	34.2	2.5	2128	2 US-08-115-593-39	Sequence 1, Appli
c	27	33.8	2.4	519	1 US-08-339-582-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 1, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA, USA
; COUNTRY: USA
; ZIP: 22213-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300 6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 89149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-F1s
; US-08-232-463-14

5-08-794-796-1 Sequence 1, Application US/08794796
 GENERAL INFORMATION:
 APPLICANT: Emery, John
 ADDRESS: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0; Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/553,619B
 FILING DATE: December 1, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Marcus-Wyner, Lynn
 REGISTRATION NUMBER: 34,869
 REFERENCE/DOCKET NUMBER: 137-1082/PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/354-3588
 TELEFAX: 415/857-1125
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2621 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 ORIGINAL SOURCE:
 ORGANISM: Chimeric tomato spotted wilt virus S RNA
 US-08-553-619B-8

RESULT 6
 US-07-867-106-2/c
 Sequence 2, Application US/07867106
 Patent No. 5389526
 GENERAL INFORMATION:
 APPLICANT: Slade, Martin B
 APPLICANT: Chang, Andy C M
 APPLICANT: Williams, Keith L
 TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
 SLIME MOULDS of the Genus Dictyostelium
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
 STREET: One Liberty Place 46th Floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/867,106
 FILING DATE: 19900625
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: AU PJ 7187
 APPLICATION NUMBER: PCT/AU90/00530
 FILING DATE: 02-NOV-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Feeney, Joanne Longo
 REGISTRATION NUMBER: 35,134
 REFERENCE/DOCKET NUMBER: RICE-0002
 COMPUTER READABLE FORM:
 Sequence 8, Application US/08553619B
 Patent No. 5919705
 GENERAL INFORMATION:
 APPLICANT: Dehaan, Petrus T.
 TITLE OF INVENTION: Virus Resistant Plants
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSE: No. 5919705arts Crop Protection
 STREET: 975 California Avenue
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94041

RESULT 5
 S-08-553-619B-8
 Sequence 9, Application US/08553619B
 Patent No. 5919705
 GENERAL INFORMATION:
 APPLICANT: Dehaan, Petrus T.
 TITLE OF INVENTION: Virus Resistant Plants
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSE: No. 5919705arts Crop Protection
 STREET: 975 California Avenue
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94041

MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: C-terminal
 ORIGINAL SOURCE:
 ORGANISM: Elmeria maxima
 STRAIN: Houghton
 DEVELOPMENTAL STAGE: sporozoite
 IMMEDIATE SOURCE:
 LIBRARY: sporozoite cDNA cloned in Lambda ZAPII
 CLONE: Em70-1
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1368
 ID: US-08-338-057-1

Query Match 2.6%; Score 35.8; DB 1; Length 1400;
 Best Local Similarity 69.0%; Pred. No. 1;
 Matches 49; Conservative 0; Mismatches 22; Indels 0; Gaps 0

Dy	1318	ttagagttatacttgtatgttttaacttggaaataattttaaaa	1377
Dy	1326	TCAAGGAAATGGTCTTAATTGTGTTAATCTGCAGGAATAATAAAAA	1385
Dy	1378	aaaaaaaaaa 1388	
Dy	1386	AAAAAAA 1396	

RESULT 9
 US-08-668-416-1
 Sequence 1, Application US/08668416
 Patent No. 5843722
 GENERAL INFORMATION:
 APPLICANT: Tomley, Fiona M.
 Dunn, Paul P. J.
 Bumstead, Janene M.
 APPLICANT: Vermeulen, Arno N.
 TITLE OF INVENTION: Coccidiosis poultry vaccine
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Akzo No. 5843722el Patent Department
 STREET: 1300 Picard Drive, Suite 206
 CITY: Rockville
 STATE: Maryland
 COUNTRY: U.S.A.
 ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentnet Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/668,416
 FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/464,164
 FILING DATE: June 2, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Gormley, Mary E.
 REGISTRATION NUMBER: 34,409
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 258-5200
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1400 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO

Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Lodes, Michael J.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Smith, John M.
 APPLICANT: McNeill, Patricia D.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF
 NUMBER OF SEQUENCES: 81
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/993,674A
 FILING DATE: 18-DEC-1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Maki, David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 210121.422C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 581 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-993-674A-22

RESULT 11
 US-08-814-306-22
 ; Sequence 22, Application US/08834306
 ; Patent No. 6054135
 GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA:
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; APPLICATION NUMBER: US/08/834,306
 ; FILING DATE: 15-APR-1997
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.422C1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 22:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 581 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-814-306-22

Query Match 2.6%; Score 35.4; DB 4; Length 581;
 Best Local Similarity 61.3%; Pred. No. 0; 83; Mismatches 0; Indels 0; Gaps 0;
 Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
 APPLICANT: HITZ, WILLIAM DEAN
 APPLICANT: Nucleotide Sequences of Canola
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF CANOLA
 TITLE OF INVENTION: AND SOYBEAN PALMITOYL-ACP THIO-
 ESTERASE GENES AND THEIR USE IN
 TITLE OF INVENTION: THE REGULATION OF FAITY ACID
 TITLE OF INVENTION: CONTENT OF THE OILS OF SOYBEAN
 TITLE OF INVENTION: AND CANOLA PLANTS
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 STREET: 1007 MARKET STREET
 CITY: WILMINGTON
 STATE: DELAWARE
 COUNTRY: USA
 ZIP: 19898
 COMPUTER READABLE FORM:

RESULT 12
 US-08-993-674A-22
 ; Sequence 22, Application US/08993674A
 ; Patent No. 6228372
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.

MEDIUM TYPE: DISKETTE, 3.50 INCH
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: MICROSOFT WINDOWS 95
 SOFTWARE: MICROSOFT WORD VERSION 7.0A
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/793,410
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10627
 FILING DATE: AUGUST 25, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: CHRISTENBURY, LYNN M.
 REFERENCE/DOCKET NUMBER: 30,971
 REFERENCE/DOCKET NUMBER: CR-9567-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-773-0164
 TELEFAX: 302-773-0164
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1303 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ANTI-SENSE: NO
 US-08-793-410-11

Query Match 2.6%; Score 35.4; DB 2; Length 3489;
 Best Local Similarity 57.8%; Pred. No. 2.1;
 Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1295 gtatcttataatattatgtataggatatacttgttatatgttttattaagct 1354
 Db 639 GTAATAATTATTTATGTATAATTATAGTATGATAATAAAAGTTACGTATCAGTT 580

Qy 1355 9tggataaaatataaaaaaaa 1387
 Db 579 ATTGGTTAGATTGACTGAAACAAAAAA 547

RESULT 14
 US-08-728-323A-1/c
 Sequence 1, Application US/08728323A
 Patent No. 5948676
 GENERAL INFORMATION:
 APPLICANT: Chang, Yuan
 APPLICANT: Bohenzky, Roy A.
 APPLICANT: Russo, James J.
 APPLICANT: Edelman, Isidore S.
 APPLICANT: Moore, Patrick S.
 TITLE OF INVENTION: Immediate Early Protein From Kaposi's
 Sarcoma-Associated Herpesvirus, DNA
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/728,323A
 FILING DATE:
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 52342
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391 0525
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 32207 base pairs
 TYPE: nucleic acid

Scoring table:	BLOSUM62	Alignments				
Searched:	100059 seqs, 36664827 residues					
Total number of hits satisfying chosen parameters:	100059					
Minimum DB seq length:	0					
Maximum DB seq length:	20000000000					
Post-processing:	Minimum Match 0% Maximum Match 100% Listing First 45 summaries					
Database :	SwissProt_39.1*					
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES						
Result No.	Score	Query	Match	Length	DB ID	Description
1	343.5	15.0	455	1	ENT1_HUMAN	Q99808 homo sapien
2	326.5	14.2	456	1	ENT2_HUMAN	Q14542 h equilibra
3	320.5	14.0	456	1	ENT1_RAT	Q04699 rattus norv.
4	301	13.1	456	1	ENT1_RAT	Q04698 rattus norv.
5	209.5	9.1	327	1	ENT2_MOUSE	Q61672 m equilibra
6	182.5	8.0	517	1	FU26_YEAST	P31381 sacccharomyces
7	134	5.8	432	1	Y672_MEJUA	Q58086 methanococcus
8	124.5	5.4	506	1	HIPB_HAELIN	P71338 haemophilus
9	117.5	5.1	559	1	YAFB_SCIP0	Q09852 schizosaccharomyces
10	113.5	4.9	477	1	Y038_MIGGE	P47344 mycoplasma
11	113.5	4.9	587	1	T9S3_MOUSE	Q9et30 mus musculus
12	113.5	4.9	589	1	T9S3_HUMAN	Q9hd45 homo sapien
13	113.5	4.9	604	1	NUSM_HORSE	P48656 equus caballus
14	113.5	4.9	506	1	TCH2_BACSU	P14512 bacillus subtilis
15	113	4.9	614	1	YDNK_LACTIC	P42377 lacticoccus
16	112	4.9	383	1	NAPA_ENTHR	P26235 enterococcus
17	111.5	4.9	397	1	PAR2_HUMAN	P55085 salmonella
18	111.5	4.9	433	1	TCR_STAU1	P02983 staphylococcus
19	111.5	4.9	439	1	UHPC_ECOLI	P09936 equus caballus
20	110.5	4.8	479	1	Y098_MYCPN	P75535 mycoplasma
21	110.5	4.8	785	1	ISP4_SCIP0	P40900 schizosaccharomyces
22	108.5	4.7	442	1	UHPC_SALTY	P27669 salmonella
23	108	4.7	388	1	YUBA_BACSU	P032086 bacillus subtilis
24	106.5	4.6	399	1	PAR2_MOUSE	P55086 mus musculus
25	106	4.6	547	1	NUSM_ASSTRU	P24884 escherichia coli
26	105.5	4.6	598	1	THIX_YEAST	P08485 sacccharomyces cerevisiae
27	105	4.6	542	1	ATRI_YEAST	P13090 sacccharomyces pombe
28	104.5	4.6	894	1	YNB6_YEAST	P27514 sacccharomyces kudriavzevii
29	104	4.5	527	1	NUSM_ACACA	P37376 acanthamoeba
30	104	4.5	552	1	NUSM_RHISA	Ozyomyces rhizophilus
31	103.5	4.5	345	1	NUOH_RHOCA	P42032 rhodobacter sphaeroides
32	103.5	4.5	606	1	NUSM_EQUIS	P2485 equus asinus
33	103	4.5	459	1	NUSM_BOVIN	P03910 bos taurus

DR	InterPro; IPR002259; DER_Nucleoside_tran.
DR	PRINTS; PF01733; Nucleoside_tran; 2.
DR	PRINTS; PRO1130; DEVENTRNSPRT.
DR	ProDom; PD005103; DER_Nucleoside_tran; 1.
KW	Glycoprotein.
FT	INIT_MET 0 0
DOMAIN	1 11
FT	TRANSEM 12 28
DOMAIN	29 81
FT	TRANSEM 82 106
FT	TRANSEM 107 110
FT	TRANSEM 111 129
FT	TRANSEM 130 137
FT	TRANSEM 138 156
FT	TRANSEM 157 173
FT	TRANSEM 174 198
FT	TRANSEM 199 205
FT	TRANSEM 206 226
FT	TRANSEM 227 290
FT	TRANSEM 291 310
FT	TRANSEM 311 322
FT	TRANSEM 323 341
FT	TRANSEM 342 358
FT	TRANSEM 359 377
FT	TRANSEM 378 392
FT	TRANSEM 393 412
FT	TRANSEM 413 430
FT	TRANSEM 431 451
FT	TRANSEM 452 455
CARBOHYD	47 47
SEQUENCE	455 AA; 50088 MW; 9098E95E26515850 CRC64;
Query Match	Score 343.5; DB: 1; Length 455;
Best Local Similarity	15.0%
Matches	24.3%; Pred. No. 7.5e-17;
	Mismatches 212; Indels 59; Gaps 13;
QY	18 PRDKYNIVVWLIVLGFGVLLPPNNMFTITTAPEYYVNVNRFKPDGV--TWSKEFMGSUT 74
DB	6 PODYKAWLIFPFMIGLGPWNFEMT-ATQYFTNRJDMSONVSLTYAELSKDAQASAA 64
QY	75 IGSQLP--NASINVNFLFLLIAGPLIYRVFA----PYCFNIVN---LTILLI 117
DB	65 PAAPLPERNSLSALFNNVMTLCAMPLLIFTYLNSFLHQRIPOSVRTLGSVLLVFLI 124
QY	118 LVVLEPTEDMSWFLYVNLGMAISNSNGLYENSIVGGDPEPTVGLLGNNCG 177
DB	125 TAILVKVQLDALP-FFVITMKIVLINSFLGQSLGLPASTAPINSQFLAG 183
QY	178 LITVVKIGTVYFLNDEPKLVAIYFGISLVILVCAFLFETKQDFYHYHQ----- 231
DB	184 FPAVAMI-CATAGSELSESAAYFITACVILITCYGLERLYRYQLKLGP 242
QY	232 -KGMEIREKAETDRPSPSLWTT----FTNCYGOLENWFCFAVT 271
DB	243 GEQETKLDLISKGEPRAKESGVSVNSQPTNEHSHTKAIKLNISLAFSVCFIFTIT 302
QY	272 LTIEPVMMTVTRDGFLINKINSENDEIYTFLVNLIAIGSIVASHIWP--TP 329
DB	303 IGMPAVVTEVKSSIAQ---SSTWERYFIPSCELTFNIDMLGRSLTAFMMPGKDS 357
QY	330 RYLFIAILRLAFIPPEFFCNCVRQTRAY-PYFFESTDFVIGGIAMSFSHGYSALAMG 388
DB	358 RWLPSLVLARLVEPPLLICN-TKPRYLTWVFEHDWAFIFMAAFAFSNGYLASLCM 415
QY	389 YTPNVVPSHYSRFAAQSLVCTLAVGLTGWLWVVIHFV 428
DB	416 FGPKVKPAEAETAGAIMAFFCLGALGAVESFLRAIV 455
RESULT	2
ID	ENT2_HUMAN STANDARD; PRT; 456 AA.

Db 67 NWWVLLSOLPLLFTLNSFLYQCIPESVRLG---SLLAIIIFALTAALVKVDLSP 122
 QY 130 SWFFWVTLGMATSINFNSGLYENSYVGDDPFHTYICALLGNINCGL---LITVVK 184
 Db 123 GLFFTSITMASWVSHFNSCAVLSQGLSFLGQGMSTYSTFLSGLAGLIAALAMTSLA 182
 QY 185 IGVTFYLADPKLVAVVF--EFTIKQ---GTSVLVIAL--- 223
 Db 183 SGV----DQPSALGYFITPCVGILISIICLSPHLKFAYYLTKKPQAVQEELETK 236
 QY 224 -DFYHYHHQKGMEIR-----EK----AETDRPSPLSATTTFTNCGOLFNV 264
 Db 237 AELLGADEKNGKIPVSPQDAGPTLDPEKEELGLEEPQKPGKPSVTVFRKIWLTLCL 296
 QY 265 WFCFAVTUTIFP-VMMTWTTRGSGLFLNKIMSENDEIYTLLTSLVLENLFAAIGSIASK 323
 Db 297 VLFVFTVLVSVEPAITAMVTTSSNS----PGKWSOFFNPICCFLLPNMDNLGRSLTSY 350
 QY 324 IHWP--TPRYLKPAIILRALFPLPPFFCNRYTRAVPFFESTDIFVGGTAMSFSHGY 381
 Db 351 FLMPDEDSQLPLLVLCRFLFLVFLMLCHVQPBAR-LPITIWFODAYFTFMFLFAISNGY 409
 QY 382 LSALAMGTYTPNVPSPHSYRFQAQLSYCTLMVGLLTG 417
 Db 410 FVSLTMCLAPRQVLPHEREVAQALMTFFLAGLUSCG 445

RESULT 4

ENTL_RAT STANDARD; PRT; 456 AA.

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 1 (EQUILIBRATIVE

DE NITROBENZYLMERCAPTOPURINE RIBOSIDE-SENSITIVE NUCLEOSIDE TRANSPORTER) (NUCLEOSIDE TRANSPORTER, NMPR-SENSITIVE NUCLEOSIDE TRANSPORTER) (NUCLEOSIDE TRANSPORTER, ES-TYPE).

DE SLC29A1 OR ENT1.

GN Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteostomi;

RA Yao S.Y.M., Ng A.M.L., Muzika W.R., Griffiths M., Cass C.E., Baldwin S.A., Young J.D.; "Molecular cloning and functional characterization of nitrobenzylthiouridine (NBMPR)-sensitive (es) and NMPR-insensitive (el) equilibrative nucleoside transporter proteins (rENT1 and rENT2) from rat tissues"; J. Biol. Chem. 272:28423-28430 (1997).

CC -!- FUNCTION: MEDIATES BOTH INFUX AND EFFLUX OF NUCLEOSIDES ACROSS THE MEMBRANE (EQUILIBRATIVE TRANSPORTER). IT IS SENSITIVE (ES) TO LOW CONCENTRATIONS OF THE INHIBITOR NITROBENZYLMERCAPTOPURINE RIBOSIDE (NBMPR) AND IS SODIUM-INDEPENDENT. IT HAS A HIGHER AFFINITY FOR ADENOSINE, RESISTANT TO DIPYRIDAMOLE AND DIAZEPAM INHIBITION (ANTICANCER CHEMOTHERAPEUTICS DRUGS).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. TISSUE SPECIFICITY: EXPRESSED IN JEJENUM, LIVER AND LUNG.

CC -!- PTM: GLYCOSYLATED.

CC -!- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.

CC -!- PTM entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	DR	EMBL; AF015304; AAC88049.1;
	DR	InterPro: IPI002229; DER_Nucleoside_tran.
	DR	Pfam: PF01733; Nucleoside_tran; 2.
	DR	PRINTS: PRO1130; DER_Nucleoside_tran; 1.
	DR	PRODOM: PD005103; DER_Nucleoside_tran; 1.
	KW	Transmembrane; Transport; Glycoprotein.
	INTR-MET	0 0 BY SIMPLASMIC (POTENTIAL)
	FT DOMAIN	1 11 CYTOPLASMIC (POTENTIAL)
	FT TRANSMEM	12 28 EXTRACELLULAR (POTENTIAL)
	FT DOMAIN	29 81 106 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	82 110 CYTOPLASMIC (POTENTIAL)
	FT DOMAIN	107 110 129 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	111 130 137 EXTRACELLULAR (POTENTIAL)
	FT DOMAIN	130 138 156 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	138 157 173 CYTOPLASMIC (POTENTIAL)
	FT DOMAIN	157 174 198 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	174 199 205 EXTRACELLULAR (POTENTIAL)
	FT DOMAIN	199 206 226 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	206 227 290 CYTOPLASMIC (POTENTIAL)
	FT DOMAIN	227 291 310 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	291 311 322 EXTRACELLULAR (POTENTIAL)
	FT DOMAIN	311 323 342 CYTOPLASMIC (POTENTIAL)
	FT CARBOHYD	343 359 N-LINKED (GLCNAC, .)
	FT TRANSMEM	360 378 393 EXTRACELLULAR (POTENTIAL)
	FT DOMAIN	379 393 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	394 413 CYTOPLASMIC (POTENTIAL)
	FT DOMAIN	414 431 EXTRACELLULAR (POTENTIAL)
	FT TRANSMEM	432 452 EXTRACELLULAR (POTENTIAL)
	FT DOMAIN	453 456 EXTRACELLULAR (POTENTIAL)
	FT CARBOHYD	453 47 N-LINKED (GLCNAC, .)
	FT DOMAIN	47 47 N-LINKED (GLCNAC, .)
	FT SEQUENCE	53 49885 MW; A34CE92C20836D9B CRC64;
	Query Match.	13.1%; Score 301; DB 1; Length 456;
	Best Local Similarity	22.3%; Pred. No. 6.2e-14;
	Matches 107; Conservative	84; Mismatches 191; Indels 98; Gaps 17;
Qy	1.8	PRDKYNIVYWLVTLVGFVLLPWNMFETIITAPEYYVNNWFKPDGV---ETWYSKEPMGS 72
Db	6	PSVSLPARISSLSAIFNNWMTLCMLPLIF---TCLNFLHOKVQSLSRLIGLSLAILLVF 122
Qy	1.6	LILVIVLPEDTDSMSWFWVTLGMATSINFNSGLYENSYVGDDFPHTYIGALLIGNNI 175
Db	123	LVTATLYVQMDALS-FFIITMIKIVLINSFGAIQOASLFGVLPANNTYAPIMSCQGL 181
Qy	1.6	CGLLITVVKI-----GVYFLINDEPKLVLAIVYFGISLVLLVCAIALFFTKQ 223
Db	182	AGFETSYVAMICAVASGSKLSESAFGYFT---ACAVV-----ILALICYLALPWN-- 228
Qy	2.4	DYFTFTNCYQOLENVWFCFAVLTIFPMMTYTRGDSGLFLNKIMSENDEIYLTLSLSPVFC 250
Db	289	WVL-----ALSVCFFITVTIGLFP--AVTAEVESSIAGTSPWKNC-YFPVACPLNF 337
Qy	3.11	NLEPAIGTSIVASKIHWI--TPRVLKFATLRAFLPIPFFFCMNYRVOTRAYPVFFESTDIF 368
Db	338	NVDWLGRSLTAICMNGQDSRKLWPLVACRVFEPILLMLCNVK-QHHYPLSFLKHWDWF 396
Qy	3.69	VIGGIANSFSHGYSALAMGTTPNVYBPHYSYFAAQSLVCTMVGLTGILGPVWVIEHFV 428
Db	397	ITFMAAFAFNSNGYLASLQMCFCGPKKVPAEAETAGNIMSFFLGGLAGAVLSFLRLAV 456

RESULT 5				
ENT2_MOUSE STANDARD; ID:ENT2_MOUSE; PRT; 327 AA.				
AC Q61672; DT 01-NOV-1997 (Rel. 35, Created) DT 01-NOV-1997 (Rel. 35, Last sequence update)				
DT 01-NOV-2001 (Rel. 40, Last annotation update)				
DE EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 2 (EQUILIBRATIVE NITROBENZYLIMERCAPTOPURINE RIBOSIDE-INSENSITIVE NUCLEOSIDE TRANSPORTER) (EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 2 (EQUILIBRATIVE DE TRANSPORTER, EI-TYPE)) (36 kDa NUCLEOLAR PROTEIN HNP36). (NUCLEOPHOBIC NUCLEOLAR PROTEIN 36 kDa) (DELAYED-EARLY RESPONSE PROTEIN 12).				
GS Mus musculus (Mouse)				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Murinae; Mus; NCBI_TaxID=10090; [1]				
RN RP SEQUENCE FROM N.A. STRAIN=BALB/C; TISSUE=Fibroblast; RXN MEDLINE=93367016; PubMed=7639753;				
RA Williams J.B.; Lanahan A.A.; RT "A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucleolar protein.";				
RL Biochem. Biophys. Res. Commun. 213:325-333(1995).				
CC -!- FUNCTION: MEDIATES EQUILIBRATIVE TRANSPORT OF PURINE AND PYRIMIDINE NUCLEOSIDES, AND THE PURINE BASE HYPOXANTHINE.				
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR. INTEGRAL MEMBRANE PROTEIN.				
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM AND A SHORT FORM/HNP36 (SHOWN HERE); SEEKS TO BE PRODUCED BY ALTERNATIVE SPLICING.				
CC -!- INDUCTION: BY PLATELET DERIVED GROWTH FACTOR (PDGF) AND FIBROBLAST GROWTH FACTOR (FGF).				
CC -!- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).				
CC DR EMBL; X86582; CA60381.1. -.				
DR MGD; MGI:1345278; SLC29a2.				
DR InterPro; IPR002259; DER_Nucleoside_tran.				
DR PRINTS; PF01733; Nucleoside_tran; 1.				
DR PRION; PR01130; DEFENTRNSPRT				
DR PRODOM; PDO05103; DER_Nucleoside_tran; 1.				
RW Nuclear protein; Transmembrane; Transport; Alternative splicing.				
FT TRANSMEM 2 22 POTENTIAL.				
FT TRANSMEM 32 52 POTENTIAL.				
FT TRANSMEM 63 83 POTENTIAL.				
FT TRANSMEM 159 179 POTENTIAL.				
FT TRANSMEM 194 214 POTENTIAL.				
FT TRANSMEM 231 251 POTENTIAL.				
FT TRANSMEM 267 287 POTENTIAL.				
FT TRANSMEM 303 323 POTENTIAL.				
SQ 327 AA; 36113 MW; 5D2D3FF4BBDD592B6 CRC64;				
Query Match 9.1%; Score 209.5; DB 1; Length 327; Best Local Similarity 22.4%; Pred. No. 8 5e-08; Matches 76; Conservative 49; Mismatches 134; Indels 81; Gaps 12;				
131 WFFWWYTLGMATSIINSFSNGLYENSYVGVGDFPHTYIGALLIGNNCGLLTVVKI----				
5 WF -----INSFCAVLQGSLFQLGTMPSYSTFLSGQLAGTAAALMLMSLAS				
186 GVTYFLNDEPKLVAVYF--GISLVLVLYCATL-----FTI-----TK 222				
Db 55 GV-----DAQTSALGYFITPCVGILLSTCYLSLPLHKPARYYLTEKLQSAPTOELTK 108				
QY 223 QDFYHYHQKGMEIR-----EKAFTDRSPSILWTTFTCYGQLFNVNE 266				
Db 109 AELOADEKNGVPISPQQASPTLDPEKEPEEPQPKPSVPEVFRKIWLTAFLCLVL 168				
QY 267 CFAVLTIFPVMM-TVTTTRGDSG --FLINKINSENDETYLTLTSFLVNLAIGSIV 320				
Db 169 VFTVPLSVFFAATANVYTSSNSPGKKGWLFNP1-----CCFLFVNMDQGRSL 217				
QY 321 ASKTHWP---TPRYLKFTTLLRFLTPFFCNYTQVQTAPVFESTDLEVIGSIAAMS 377				
Db 218 TSYFWPDESSQQLPLVLCRLEPVFLCHVPOHAR LPPIFRQDAYFITFMLFAV 276				
QY 378 SHGYSALANGYTPNVSYPHSYSRFAAQSLSYCTLMGLLTG 417				
Db 277 SNGYLVLSTMCLAPRQLPHEREVAAGALMTFFLAGLGS CG 316				
RESULT 6				
FU26 YEAST ID FU26 YEAST STANDARD; PRT; 517 AA.				
AC P31381; DT 01-JUL-1993 (Rel. 26, Created)				
DT 01-JUL-1993 (Rel. 26, Last sequence update)				
DT 20-AUG-2001 (Rel. 40, Last annotation update)				
DE NUCLEOSIDE TRANSPORTER FUN26.				
GN FUN26 OR YAL022C.				
OS Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomyces cerevisiae (Baker's yeast).				
OC Saccharomyces cerevisiae; Saccharomycetales; Saccharomyces.				
OX NCBI_TaxID=4932; RN [1]				
RP SEQUENCE FROM N.A.				
RC MEDLINE=S288C / AB972;				
RC MEDLINE=9320952 / PubMed=8458570;				
RA Ouellette B.F.F.; Clark M.W.; Keng T.; Storms R.K.; Zhong W.W.; Zeng B.; Fortin N.; Delaney S.; Barton A.B.; Kaback D.B.; Bussey H.; RT "Sequencing of chromosome I from <i>Saccharomyces cerevisiae</i> : analysis of a 32 kb region between the LTEL and SPOT genes.";				
RL Genome 36:32-42(1993).				
RN [2]				
RP CHARACTERIZATION				
RX MEDLINE=20408929; PubMed=10827169;				
RA Vickers M.F.; Yao S.Y.; Baldwin S.A.; Young J.D.; Cass C.E.; RT "Nucleoside transporter proteins of <i>Saccharomyces cerevisiae</i> . Demonstration of a transporter (FU1) with high uridine selectivity in plasma membranes and a transporter (FUN26) with broad nucleoside selectivity in intracellular membranes.";				
RL J. Biol. Chem. 275:25931-25938(2000).				
CC -!- FUNCTION: HAS BROAD NUCLEOSIDE SELECTIVITY (URIDINE, ADENOSINE AND CYTIDINE) AND MOST LIKELY FUNCTIONS TO TRANSPORT NUCLEOSIDES ACROSS INTRACELLULAR MEMBRANES.				
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).				
CC -!- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC DR L05146; AAC04935.1; -.				
DR L05027; AAA7039.1; -.				
DR S36712; S36712.				
DR S0000020; FUN26.				
DR InterPro; IPR002259; DER_Nucleoside_tran.				
DR Pfam; PF01733; Nucleoside_tran.				
DR ProDom; PDO05103; DER_Nucleoside_tran; 1.				
KW Transmembrane; Transport.				
FT TRANSMEM 76 96 POTENTIAL.				

Query Match 8.0%
 Best Local Similarity 19.9%
 Matches 79; Conservative 69; Mismatches 131; Indels 117; Gaps 20;

Query 23 NIYVWLVLVGFVLLPQPNMFTIATPEYYVNWFQKPDGVETW--YSKEFMSLTTIGSQL 79
 Db 75 NLSYITFAAIGIGLWLWNICL-SASQYFKHDFKDTSI-WAKIFTSSMMMSFTISSL 1.31

Query 80 PNASINVNLFL-----TIAQPLIYR-VFAPICFNIVNLNTILIVTLEPT 1.27
 Db 132 -----ENIYLAKRQKYSSRRVING-LWWEITVFTYMCF----FTLHLFL----- 1.71

Query 128 SMSGWFWVTLGMATSII-NFSNGLYENSYYVGQGDEPHTYIGALLIGNICGLITYVKIG 1.86
 Db 172 -PWENFMFIMLWVVISMGMTAQNGIMAIAVFGSEYSGQMVCGAVAGLPLSLVLA 2.30

Query 187 VTVFLNDEPKLVA-IVFGISLVLLVCAAILFFTK----- 2.22
 Db 231 LAFIENSSVSTGGILLYFTTLLVVIC-VVMSVSKRSRKVYNENNVEDGHITDVLLG 2.89

Query 223 -----QDFTHYH-----HQKGMAIREKAETDRPSPLSLWTTNC 2.57
 Db 290 SLSNEEIRYGRIDQMEDEHARTNGTRDNDGEPLQLKVPF-----VLFPAKL 3.41

Query 258 YGOLFNYWFCFAVTLTIPVMUTVTQDSGFLNKIMSENDELYLTTSFLVNLFAAG 3.17
 Db 342 KYLVLISIFTTQPLPSNAQYIPLI--FTLWNLGDLYG 3.92

Query 318 SIYASKIHW-----TPRYLKFAILRLALEPFF 3.46
 Db 393 RVIAD--WPMFRDQKETPKTFIYSURVAALPLF 4.25

RESULT 7
 Y672_MEJUA STANDARD PRT; 432 AA.
 ID Y672_MEJUA
 AC Q56086;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN MJ0672.
 GN
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 NCBI_TAXID=2190;
 OX RN
 RP SEQUENCE FROM N.A.
 STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; Published=8888087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.C., Blake J.A., Fitzgerald L.M., Clayton R.A., Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Overbeek R., Kirnahan E.F., Weinstock K.G., Merrick J.M., Scott J.L., Geoghan N.S.M., Weidman J.F., Fuhrmann J.L., Utterback T.R., Kelley J.M., Peterson J.D., Saclow P.W., Hanna M.C., Cohen M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Ventor J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.",

RL Science 273:1058-1073(1996).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE NADC/P/PHO87 FAMILY OF TRANSPORTERS.
 CC NADC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the FMBI outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC EMBL: U67514; AAB98666; 1; -.
 DR TIGR: M0672; -.
 DR InterPro: IPR001898; Na_sulph_symp.
 DR Pfam: PF00939; Na_sulph_symp; 2.
 DR PROSITE: PS01271; NA_SULFATE; 1.
 KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
 FT TRANSMEM 7 27
 FT TRANSMEM 7 27
 FT TRANSMEM 29 49
 FT TRANSMEM 68 88
 FT TRANSMEM 1.24 1.44
 FT TRANSMEM 1.56 1.76
 FT TRANSMEM 1.96 2.16
 FT TRANSMEM 2.41 2.66
 FT TRANSMEM 3.21 3.46
 FT TRANSMEM 3.58 3.78
 FT TRANSMEM 3.79 3.99
 FT TRANSMEM 4.12 4.32
 SQ SEQUENCE 432 AA; 4/541 MW; A24BF964475CB74 CRC64; ;

Query Match 5.8%; Score 134; DB 1; Length 4.32;
 Best Local Similarity 19.1%; Pred. No. 0.017; Mismatches 80; Gaps 22; Matches 82; Conservative 133; Indels 134;

Query 66 SKEEFMPG-----SLTIGQQLPNSINVNLFTIAGPLIYRFAPVCNIVNTILLV 1.19
 Db 4 SKEEFIGLITIASLIFGSSLPLDPYKG1-VILIVAGCWWFEFLPLPVTSLAIPIMAWFL 6.1

Query 120 IVLEPTEDSMSNP-----FWVTLG-----MATSNFSN-----GLYENSVYGVGDGPHT----- 1.64
 Db 62 GTFN-LKALTYFAHPFLFPLGGFMLAQALKHNLDKFIAKLLNGKDFKTTFCFLMFL 1.20

Query 165 - YIGALIIGNNIGLILTVVKIGTVY----- 1.64
 Db 121 SAYFLSMWISNTSATLPLTALGILHKTKGKLRFDFLLGVAYASASIGGIATIGSPNA 1.80

Query 198 VALVY-----FGTQDFHYHQQGMETREKA 2.40
 Db 181 IASSYLDYGGFSWFKVGPISLFLCLCTLYIYFKWIKPED-----IAIQARM 2.31

Query 241 EDTRPSPSI-----LWTTFTNCYQGLFNWFCFAVTLTIFPVYMMTVTRGDSFLN 2.91
 Db 232 ELRNAYKLLVTVLTLASLW-TISDYLEIFNVQY-DSVTAIAFLIL-----FVF 2.81

Query 292 KIMSEND----ELYLL-----TSFLVENLFAAIGSVASKIHWPTPR 3.30
 Db 282 NLVEVNDFKKIKWGTLLFGGALCLGGVIVKSGANTPLSEKUJIALGNL-----TPI 3.33

Query 331 YLFKAIILRALEPFPFFCNYRVQTRAYPVFEST-----DIFVIGIAMSPTSHGQLSAL 3.85
 Db 334 VLLFLVTTITLIL-NETTSNTGLTGIVPILEFGVSLGIPKEIIL-AVGMASCSFTI-L 3.89

Query 386 AMGYTPNNV 3.94
 Db 390 PVGTPPNV 3.98

HITB_HAEIN STANDARD; PRT; 506 AA.
ID HITB_HAEIN; Q33440;
AC P71339; 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE IRON(III)-TRANSPORT SYSTEM PERMEASE PROTEIN HITB.
GN HITB OR H1098
OS Haemophilus influenzae.
OC Haemophilus; Proteobacteria; gamma subdivision; Pasteurellaceae;
NCBI_TAXID=727;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=2542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., FitzHugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shizley R., Liu L.T., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom B., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geohagen N.S.M.,
RA Gnehn C.L., McDonald L.A., Fraser C.M., Small K.V., Fraser C.M.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.";
RT Science 269:496-512(1995);
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NTHI TN106;
RA Sanders J.D., Cope L.D., Hansen E.J.;
RX MEDLINE=95012644; PubMed=7927717;
RA RT "Identification of a locus involved in the utilization of iron by
Haemophilus influenzae".
RL Infect Immun. 62:4515-4515(1994).
CC -!- FUNCTION: INVOLVED IN A PEROPLASMIC BINDING-PROTEIN-DEPENDENT
CC IRON(III) TRANSPORT SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -!- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE CYSTW
CC SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC DR EMBL; U32695; AAC21774_1; ALT_INIT.
DR EMBL; S72674; ABB32111_1;
DR TIGR; HI0098;
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp_2.
DR PROSTRE; PS00405; BPD_TRANSPIINN_MEMBR; FALSE NEG.
KW Transport; Iron transport; Transmembrane; Inner membrane;
KW Complete proteome.

FT TRANSEM 9 29 POTENTIAL.
FT TRANSEM 57 77 POTENTIAL.
FT TRANSEM 90 110 POTENTIAL.
FT TRANSEM 116 136 POTENTIAL.
FT TRANSEM 174 194 POTENTIAL.
FT TRANSEM 218 238 POTENTIAL.
FT TRANSEM 275 295 POTENTIAL.
FT TRANSEM 314 334 POTENTIAL.
FT TRANSEM 350 370 POTENTIAL.
FT TRANSEM 379 399 POTENTIAL.
FT TRANSEM 428 448 POTENTIAL.
FT TRANSEM 480 500 POTENTIAL.
FT VARIANT 72 72 F -> L (IN STRAIN TN106).

FT VARIANT 104 104 G -> C (IN STRAIN TN106).
FT VARIANT 116 116 V -> G (IN STRAIN TN106).
FT VARIANT 167 167 S -> F (IN STRAIN TN106).
FT VARIANT 252 252 T -> M (IN STRAIN TN106).
SQ 506 AA; 56546 MW; ADA28861C1481AID CRC64;

Query Match 5.4%; Score 124.5%; DB 1; Length 506;
Best Local Similarity 21.6%; Mismatches 144; Indels 167; Gaps 27;
Matches 105; Conservative 69; Matches 144; Pred. No. 0.088;

Qy 26 YW---LVLVGFVLLPWNNMFTIAPEY---YVNYWKFDPGVETWSKE-FMGSLTIGS 77
Db 7 FWLTLILIGLPLCLPFLYILRATGEGLTRSVELLFRPMAELLSNTMLLVCVIGA 66

Qy 78 QLPNASTINVENFLIFLIAQPLIYRVAPPVCIVNLUTILLIVLEPTEDMSW---FFW 134
Db 67 ---ISGTFCAFLER---YRFEGKAFFEVAMTLPCLCPAFVSGFTWLSTFRVEFW 118

Qy 135 VTLMATMSINF-----SNGLYEENSYVGCGDPEPHYIGAL----LIGNN 174
Db 119 GTIGIMTLLSSPLAYLPVSALKRDRSLEEVSL-SLGKSPVYTAISPOLKAIGSS 177

Qy 175 ICGLLIT---VVKIGVYTFLNDEPKLVAI----VYFG-----ISLVILLVCAI---A 216
Db 178 I-LLIATHMLVLFGAWSILNYQTFTTAIFQBEYNSFNNSTAAALLSAVLMACILIVFGE 235

Qy 217 LFEPITRQDFHYHHQKGMEIREKAETDRPSSTILWTFTNCYQQ-----LFNVWFCFAV 270
Db 236 IFFRGQTLP---YHSGRKV-----TRP-----YLVKTLSFGKOCLTGFGEFSSTIFLSI 280

Qy 271 TLTIIFPMMM-----TVTTTRGD---SGFLNK-IMSENDEIYLTTSF-----307
Db 281 GV---PVTMILYWLTVGTSLESAGDSSELSAFASNFELISGIGALTVMCALPLVWAVER 337

Qy 308 -----LVENLFAAIG-----SIVASKTHWPTPRYKFAITLRAFT-----343

Db 338 YRSYLTIWIDRPLYLHAPGLVIALSIVYPSIHYANDLYOTFFVIIAYFMLYLPMAQT 397

Qy 344 -----PFIFFCNRYVROTPAYPAFPFFESTDIFVIGGIANSFHSG 380
Db 398 TURASLSDQIEKVQGSTLRNPFFYF-----RTLTPA-----ILPGVAAAFAFALV 444

Qy 381 YLSAL 385
Db 445 FLNLN 449

RESULT 9
YAEC_SCHPO STANDARD; PRT; 559 AA.
ID YAEC_SCHPO AC 009852;
AC 009852;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE INORGANIC PHOSPHATE TRANSPORTER C23D3.12.
GN SPAC2.D3.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomyces; Schizosaccharomyces.
OC Schizosaccharomyces; Schizosaccharomyces.
OX SCHIZOSACCHAROMYCETACEAE; SCHIZOSACCHAROMYCETACEAE;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Niblett D., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT 1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HIGH-AFFINITY TRANSPORTER FOR EXTERNAL INORGANIC
CC PHOSPHATE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. STRONG, TO
CC YEAST PHO84.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Informatics Institute and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	DT	20-AUG-2001 (Rel. 40, last annotation update)
CC	DE	HYPOTHETICAL PROTEIN MG098.
CC	GN	MG098.
CC	OS	Mycoplasma genitalium.
CC	OC	Bacterium; Firmicutes; Bacillus/Clostridium group; Mollicutes;
CC	OC	Mycoplasmataeae; Mycoplasma.
CC	NCBI_TaxID	2097;
CC	OX	
DR	RN	SEQUENCE FROM N.A.
DR	RP	STRAIN=ATCC 33530 / G-37;
DR	RX	MEDLINE=16026446; PubMed=7569993;
DR	RA	Froese C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Fritchman J.L., Weidman J.F., Smalkin K.V., Sandusky M., Fuhrmann J.L., Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M., Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S., Peterson S.N., Smith H.O., Hutchinson C.A. III, Venter J.C.; RT "The minimal gene complement of Mycoplasma genitalium.";
DR	RA	Science 270:397-403 (1995).
DR	RN	[2]
DR	RP	SEQUENCE OF 46-155 AND 278-382 FROM N.A.
DR	RC	STRAIN=ATCC 33530 / G-37;
DR	RX	MEDLINE=94075530; PubMed=8253680;
DR	RA	Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III; RT "A survey of the Mycoplasma genitalium genome by using random sequencing."
DR	RL	J. Bacteriol. 175:7918-7930 (1993).
DR	CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
DR	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	DR	EMBL: U01701; AAC71316; 1;
DR	DR	EMBL: U01701; AAB01013; 1;
DR	DR	TIGR: MG098;
DR	DR	InterPro: IPR003837; Glu-trNAcIn;
DR	PFAM: PF02686; Glu-trNAcIn; 1.	
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:	MG098;
DR	InterPro:	IPR003837; Glu-trNAcIn;
DR	PFAM:	PF02686; Glu-trNAcIn; 1.
DR	KW	Hypothetical protein; Transmembrane; Complete proteome.
DR	PT	TRANSMEM . 23 . 43 . POTENTIAL.
DR	EMBL:	U01782; AAD12771; 1;
DR	TIGR:</td	

Oy	180	-----ITVVKIG-----VTVYFL-----NDE-----PK	196	Query Match Score 4.9%; Best Local Similarity 10.2%; Pred. No. 0.58; Mismatches 127; Indels 103; Gaps 14;
Db	164	KTLSDFAYLVLVGMAFSELTIPKQDFYHQQGMETRKAENDRPSPLTNTLNSLQR	223	Matches 62; Conservative 49;
Oy	197	LVAIVYFGISVTLVCAIALETFITPKQDFYHQQGMETRKAENDRPSPLTNTLNSLQR	256	Qy 96 PLIYRVFAPCFVNIVNLTTILIVLVEPTEDSMSKRFEWYTLGMASTINFNSGLYENSYV 155 Db 286 PLIFSSLIGSGCQIFAVSLIVAMETDLYTERGMSLTAIFVTAATSPVNGYFGGSLY 345
Db	224	YVQIIFYFAMWVITIPMVLYWVILQKHKNAYASKFERRYKHANHQFTLFVL--TKEN	281	Qy 156 GVGGD --FPHTYIGALLIGNICGLITVVKIGVTVYFLNDEPKLVAVY ----FGI 205 Db 346 ARQGGRRWKQMFAGLIPAMVC -----GTAFFIN ---FLAIYHHSRAIPFG- 391
Oy	257	CYGOLOPNWPGCAVLTITPVYMMTTRGDSEGLNLKIMSENDEYLTTSFLVNFNFAAI	316	Qy 206 SLVLLIVCATALEFFITQDFY -----HYHQKGMEERAKETDRPSLIVW---TTFTN 256 Db 392 -TMVAVCCIFEVILPVLNLGTILGRNLSCQPNFQLRNLSPRDSGFNLKIMSEN 297
Db	282	WFLYLINV-----ITLAT-----TSLMLMINI--AF	304	Qy 257 CYGQL -----FNUWFCFAVTLTIFPVMMTVTGTDGSFLNLKIMSEN 297 Db 449 CLGGILFGSFIEMYFLIFTSEWAYKYYYYGEMMILVVLICIVTY-----494
Oy	317	GSIIVASKIHWPTPRYLFALLRLAFLIPFFECNYRVQTRAYPVFESTDIEVIG-----	371	Qy 298 DEIYLTLSFLVNLPAIAGSIVASKIHWPPRYLKFAILRLAFLIPFFECNYRVQTRA 357 Db 495 -CVTIVCTYFLN -----AEYDRWQNTSFLSAA---STAIYVYMSYYYYFFKTKM 541
Db	305	IPFDFIQ--TTGOTYDFEWLVLRLFPLFLDILV--TYPILLLTPMLRGFKTVAA	358	Qy 358 YPVFFFESTDIFVIGGIAMSFSHGILSLS-----AMGY 389 Db 542 YGLF-----QTSFYFGYMAWFSTALGIMCGAIGY 570
Oy	372	-----GIAMSFSHGVLSAALMAGTPNVPVDSH	397	RESULT 12
Db	359	SETQTRGKIKSFSFD--MQSLIM--PNVI-SH	384	T9S3_HUMAN STANDARD: PRT: 589 AA. TD T9S3_HUMAN STANDARD: PRT: 589 AA. AC Q9HD45; Q9UHW8; Q9NWLB; Q9PGQ9; RN [1] RP [1] SEQUENCE FROM N.A. RA Warner S.J., Lonax M.I.; RA "Evolution of the TM9 super family of membrane spanning proteins."; RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases. RL "SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE)." RA [2] RN SEQUENCE OF 79-587 FROM N.A. RA Straubberg R.; RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases. CC "- SIMILARITY: BELONGS TO THE NONASPARTANIN (TM9SF) FAMILY." CC "This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)." CC DR EMBL; BC004799; AAF98160_1; RW SIGNAL; Transmembrane. FT CHAIN 1 26 POTENTIAL. FT CHAIN 27 587 TRANSMEM 9 SUPERFAMILY PROTEIN MEMBER 3. FT TRANSMEM 222 242 POTENTIAL. FT TRANSMEM 292 312 POTENTIAL. FT TRANSMEM 326 346 POTENTIAL. FT TRANSMEM 358 378 POTENTIAL. FT TRANSMEM 387 407 POTENTIAL. FT TRANSMEM 447 467 POTENTIAL. FT TRANSMEM 480 500 POTENTIAL. FT TRANSMEM 517 537 POTENTIAL. FT TRANSMEM 549 569 POTENTIAL. FT CARBOHYD 172 172 N-LINKED (GLCNAC . .) (POTENTIAL). FT CARBOHYD 417 417 N-LINKED (GLCNAC . .) (POTENTIAL). SQ SEQUENCE 587 AA: 67544 MW: 22FD4.F8588FEC2AB CRC64;

IN POSITIONS 184; 198; 370 AND 446.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

CC DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
 CC OS MTND5 OR ND5.
 CC OS Equus caballus (Horse).
 CC OG Mitochondrion.
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 CC NCBI_TaxID=9796;
 RN RN
 RP SEQUENCE FROM N.A.
 MEDLINE=95047450; PubMed=7958896;
 RX RX
 RA Xu X.; Arnason U.;
 DR "The complete mitochondrial DNA sequence of the horse, Equus caballus: extensive heteroplasmy of the control region.",
 RT RT
 RT Gene 148:357-362 (1994).
 CC !- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).
 CC ---
 CC DR EMBL: X79547; CAA6089_1;
 DR InterPro: IPR003916; NADHub_oxidredctse5;
 DR InterPro: IPR001750; Oxidored_q1_N.
 DR InterPro: IPR001516; Oxidored_q1.
 DR Pfam: PF00361; oxidored_q1_1.
 DR Pfam: PF00662; oxidored_q1_N_1.
 KW KW Oxidoreductase; NAD; UbiQuinone; Mitochondrion.
 SQ SEQUENCE 604 AA; 68056 MW; 5E794F1730235C10 CRC66;
 Query Match 4.9%; Score 113.5; DB 1; Length 604;
 Best Local Similarity 19.9%; Pred. No. 0.59;
 Matches 122; Conservative 74; Mismatches 181; Indels 237; Gaps 33;

Qy 2 VIINRSNTAYVEQAFAFRDKYNIIVYLWILVYLGFGVLLPWNMFITIAPYYVNYW---- 55
 Db 17 IMISIIFTY-KNSTPFHVKTNTSYAFT----SLTFITMMEHSQFETIISNWHMTMQ 70
 Qy 96 PIYRVEAPVCNFVNVLVILVILEPTEDMSWFWTVLGMATSINFNSGLYENSVY 155
 Db 288 PLFSSLIGSGCQIFAVSLIVIIVAMIEDLYTERGSMLSATAFYVAATSPVNGYFGSLY 347
 Qy 156 GVGDD----FPHTYTGALLGNIGNCLLITVVKIGTVLFLNDEPKLVY----FGI 205
 Db 348 ARGGRRWKIQMPQIFALPAWC----GTAFIAIN----FIATYYHHSRAIPFG- 393
 Qy 206 SLVLLVCAIALEFFITKQDFY----HYHQKGMEIREKAETDRPSPLW---TTFTN 256
 Db 394 -TMVAVCCICFFVILPILNLGVTLGRNLSSGFNFPCRVNA-VPRPIPEKKMPEAV 450
 Qy 257 CYGOL----FNVWFCFAVLTLLFPVMMTWTGDSGLNKMSEN 297
 Db 451 CQGILPFGSIIEMYFIFTSWAYK1YYVGFMMVLVILCIVT----- 496
 Qy 298 DIELTYLTSFLYNNLFAAIGSTVASKHWTPTGFLPATILAFTEFFENRYVOTRA 357
 Db 497 -CVTIVCTYFLIN----AEDYRWQWTSFLSA--STAIYVIMYSFVYYFFRTKTM 543
 Qy 358 YPVFFESTDIFVIGGIAMSFSHGYSLSA-----AMGY 389
 Db 544 YGLF-----QTSFYGYMAVFASTALGIMCGAIGY 572
 Qy 196 KLYA-----IVYFGIS--LVLYLCAIALF---FITKQDFYHYHOKGMETRE 238
 Db 178 IGFIMAMPWFLENNTNDQQIFMLDPNLTNLPLGLLAAAGKSQAFGLHPWLPSAMEG 237
 Qy 151 -----ENSYVGDDFPHYIGALIGN----NIC-GILLITYKIGTYFLNDEP 195
 Db 238 PTPVSALHSSTMNVAGVFLLRPHMENNTQSLTLCAGNTLTCAITALQNDIK 297
 Qy 104 PVCFNIVNUTIL-----ILVIVLEPTED 127
 Db 118 YLTFLITMMILVANNFLQFLIGWEGYGMFSLLIGWYGRDTANTALQALYNRGD 177
 Db 71 TUKLSSLTKLD----YPMIF-----VPVALFVWSMEFSLWYMHSDPYITREFK 117
 Qy 128 -----SMWFFFWTT-----LGM--ATSINFNSGLY---- 150
 Db 298 KIAFSTSSQSLGMIVTGINOPYLAFLHICTHAFFKAMLFCGSSTHSLNDE-QDTRK 356
 Qy 239 KAETDRPSLPSIWMTTFTNCYQL-----FNVWFCFAVLT 273
 Db 357 MGGLFNAMP----FTTTSUILLGSHALTGIPFLTGFSKDLIETANTSYNAW--ALMT 410
 Qy 274 IFVMMMTV-TTR----GDSGF--LINKMSENDEYLTLLTSFLYNNLFAAIGSIVAS 322
 Db 411 LTTASLPAVSYRLLPFLGQPRFLTSINNNPLINSKRLIGSIFA-GFFISN 468
 Qy 323 KIHWPPT----PRYLKFAILRLAIPFFFEC-----NYVYQTRAYPVFFEST 365
 Db 469 NI-YPTVPEMTMPYMK---LTALAVTILGFTLABLESLMTHNLKE-----HST 515

RESULT 1.3
 NUM_HORSE STANDARD; PRT; 604 AA.
 ID NUM_HORSE
 AC P48656;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)

Qy	366 DIFVIGGIAMSFSHGYSALAMGTYTPVVPSPHSYSRFAAOLSVCTLMVGILTTGGLW-----	420	Db	13 YSVLEFWLCLISFFSVLN - PMVLNVSLPDIANHNTTGPITNWNTAYMLTFSIGTAVVY 70
Db	516 NVF -----KFSN -----LLGYPTIM - HRLPPLANLMSOKSASLLSD1WLEN1 560		Qy	82 ASINVFN - FLLIAGPLLYRVFAPCFNIVNLILLIVYLEPTEDMSWEFWNTLGMA 140
Qy	421 PVVIEHFVDKPSIL 434		Db	71 KLSDYINIKRKLITG ----- ISLSCLSLIAFI ----- GHNHEFFLIFG -- 109
Db	561 PKSISQFQMKTSIL 574		Qy	141 TSINFSGNLXENSIVVG - GDFPHT ----- YIGALL ----- IGNNI 175
RESULT 14				
TCR2_BACSU	STANDARD;	PRT;	459 AA.	
ID P14512;				
AC P14512;				
DT 01-JAN-1990 (Rel. 13, Created)				
DT 01-OCT-1990 (Rel. 13, Last sequence update)				
DT 01-OCT-1994 (Rel. 30, Last annotation update)				
DE TETRACYCLINE RESISTANCE PROTEIN.				
GN TET.				
OS Bacillus subtilis.				
OG Plasmid pNS1.				
OC Bacteriae; Firmicutes; Bacillales/Clostridium group;				
OC Bacillales/Staphylococcus group; Bacillus.				
OX NCBI_TaxID=1423;				
RN [1]				
RP SEQUENCE FROM N.A.				
RA Nozuchi N., Aoki T., Sasatsu M., Kono M., Shishido K., Ando T.;				
RT "Determination of the complete nucleotide sequence of pNS1, a staphylococcal tetracycline-resistance plasmid propagated in <i>Bacillus subtilis</i> ."				
RT FEMS Microbiol. Lett. 37:283-288 (1986).				
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIporter.				
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).				
CC EMBL; M61217; AAA19179_1; -.				
DR PIR; S42238; S42238.				
DR InterPro; IPR001411; TCR_TetB.				
DR PRINS; PR01036; TCRTEB.				
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.				
FT TRANSEM 12 33 POTENTIAL.				
FT TRANSEM 81 100 POTENTIAL.				
FT TRANSEM 111 129 POTENTIAL.				
FT TRANSEM 140 162 POTENTIAL.				
FT TRANSEM 165 185 POTENTIAL.				
FT TRANSEM 201 221 POTENTIAL.				
FT TRANSEM 223 240 POTENTIAL.				
FT TRANSEM 256 276 POTENTIAL.				
FT TRANSEM 297 317 POTENTIAL.				
FT TRANSEM 324 344 POTENTIAL.				
FT TRANSEM 346 365 POTENTIAL.				
FT TRANSEM 432 451 POTENTIAL.				
SEQUENCE 459 AA; 50695 MW; FE33C55335372B41 CRC64;				
Query Match 4.9%; Score 113; DB 1; Length 459;				
Best Local Similarity 18.1%; Pred. No. 0.49;				
Matches 82; Conservative 79; Mismatches 133; Indels 160; Gaps 20;				
Qy 22 YNIYWWLVLVGFYLLPWNMFITTAPEYYNNWFKPDGVETWWSKEFGSLTIGSQLPN 81				
Query Match 4.9%; Score 113; DB 1; Length 614;				
Best Local Similarity 19.3%; Pred. No. 0.65;				
Matches 98; Conservative 66; Mismatches 149; Indels 196; Gaps 26;				
Qy 21 KYNITYWWLVLVGFYLLPWNMFITTAPEYYNNWFKPDGVETWWSKEFGSLTIGSQLPN 81				
Db 16 KYNFEEYIIIL-----FIFTIATEPFYRGNFHAGNDFAFNRYVMSTISALKDG 64				

Qy 55 ---WFKPDGV-----TWYSKKEPMGSLTIGSQLPNASTINVFNLFLIAAGPLIYRVFAPVC 1.06
 Db 65 OVIPOFDNALSGFGYAN---NFEYG-----PIAPTYFISVTK-FIVWSLISFSLEFSLC 1.15
 Qy 107 FNIVNLTILILVIVLEPTEDSHSWEFNVTLOGATSFNSNGLYENSVYGGDFEPHTYI 1.66
 Db 116 LFISGIFFFNESSFLRDHTNS-KLFGLLAVALEFT--FSNSTYINLYYAAPSQPLALL 1.71
 Qy 167 GALLI--GNN-----ICG-----LJITVVKIGV 1.88
 Db 172 FVILLFQRMNKMFENKRSPAAFLMVAFGAAGLPLSHTVTTCTLPFVLLYLRLIRKG-- 2.29
 Qy 189 YFLNDEPKLVAVVYFGISLVLVCALEF-----ITKQDFYH----- 2.27
 Db 230 -NPKENKIGIGFLSYSAI---GISAFFLPLLENLKSGTYNNNSNDSRSFGNNNI 2.84
 Qy 228 -YFHQKGMEIRKAETDPSPLSILWTF-----TNC-YGOLFNVWFCFAVT 2.71
 Db 285 AYFOGKNEPLYKTEFSYKFPSSLFLVIVLFLFISLINFKTNAYSLIFS--CFSLV 3.41
 Qy 272 LTIFPVAMTVTRGDSGRFLNKIMSENDEIYTLLTS---FLVFNLAIG----SIVAS 3.22
 Db 342 LIVLMQLPI-----FPWKLES---IFTVQDPARFSTIQLFSALSLSVLILPILD 3.88
 Qy 323 KTHWPTPRYLKAIIIRALFIPPEFFCNRYVOTRAYFV-----FESTDIF 3.68
 Db 389 KSGKTSVYLIGLVLVISLAEFVN-RICKGSQPLFASQSLNKTTPNMYMPDSI 4.47
 Qy 369 VIGGIAMSFSHGYLSSALAMGYSPNVVPSH 3.97
 Db 448 AIG-----EYLUQVIGSH 4.60

Search completed: February 27, 2002, 17:02:49
 Job time: 195 sec

protein - protein search, using sw model						
on:		February 27, 2002, 16:56:54 ; Search time 17.49 Seconds (without alignments) 1890.209 Million cell updates/second				
Title: US-08-816-001E-63						
Effect score: 2294						
Sequence: 1 MVIINRNSNTAYVEQPAFPRD.....LTGGGLMPVYIEHFVYDKPSSL 434						
Scoring table: BLOSUM62						
Gappen: Gapop 10.0 , Gapext 0.5						
Searched: 219241 seqs, 76174552 residues						219241
Total number of hits satisfying chosen parameters:						
Minimum DB seq length: 0						
Maximum DB seq length: 2000000000						
Post-processing: Minimum Match 0%						
Maximum Match 100%						
Listing first 45 summaries						
Database :						
1: pir1:*						
2: pir2:*						
3: pir3:*						
4: pir4:*						
Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
Result No.	Score	Query Length	DB ID	Description		
1	2265	98.7	434	2 T16080	hypothetical prote	
2	716.5	31.2	461	2 T28006	hypothetical prote	
3	706.5	30.8	450	2 T23528	hypothetical prote	
4	574.5	25.0	800	2 T23247	hypothetical prote	
5	313.5	13.7	428	2 T01489	hypothetical prote	
6	249.5	10.9	408	2 C86156	hypothetical prote	
7	214	9.3	326	2 JC4196	36K hydrophobic nu	
8	209.5	9.1	327	2 JC4195	36K hydrophobic nu	
9	196.5	8.6	513	2 T21887	hypothetical prote	
10	182.5	8.0	517	2 S36712	FUN26 protein - ye	
11	180	7.8	143	2 T22164	hypothetical prote	
12	144	6.3	418	2 D85054	hypothetical prote	
13	138.5	6.0	418	2 B85054	hypothetical prote	
14	134	5.8	432	2 H64383	Na+ transporter -	
15	128.5	5.6	382	2 G96641	hypothetical prote	
16	126	5.5	418	2 F39836	transporter BH2694	
17	125.5	5.5	674	2 T21217	hypothetical prote	
18	117.5	5.1	505	1 D64048	iron (III) ABC tra	
19	117.5	5.1	559	2 S62503	inorganic phosphat	
20	115	5.0	461	2 B83601	probable transport	
21	114.5	5.0	449	2 C75053	DNA damage-inducib	
22	113.5	4.9	477	2 H64210	eggshell protein P	
23	113.5	4.9	604	2 T11867	NADH dehydrogenase	
24	113	4.9	459	2 S42238	tetracyclin resist	
25	113	4.9	614	2 S40085	hypothetical prote	
26	112.5	4.9	694	2 E69143	hypothetical prote	
27	112	4.9	383	2 A42111	Na+/H+ -exchanging	
28	111.5	4.9	388	2 S66519	protease-activat	
29	111.5	4.9	397	2 S66518	Na+/H+ -exchanging	

30	111.5	4.9	440	1	RGBEUC
31	111.5	4.9	497	2	S6834
32	111	4.8	494	2	A82294
33	110.5	4.8	419	2	A70814
34	110.5	4.8	479	2	S71921
35	110.5	4.8	776	2	S41495
36	110	4.8	394	2	C85164
37	110	4.8	444	2	G70042
38	109.5	4.8	1152	2	E71667
39	109	4.8	480	2	B70367
40	108.5	4.7	442	2	C41853
41	108.5	4.7	507	2	T47021
42	108	4.7	507	1	D70006
43	107.5	4.7	523	2	T11317
44	107	4.7	401	2	F73037
45	107	4.7	1854	2	T39035

ALIGNMENT S

RESULT	1
T16H00	hypothetical protein F16H11.3 - Caenorhabditis elegans
C;Species:	Caenorhabditis elegans
C;Date:	20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
R;WU_X	
submitted to the EMBL Data Library, April 1996	
A;Description: The sequence of C. elegans cosmid F16H11.	
A;Accession: T16080	
A;Status: preliminary; translated from GB/EMBL/DBJ	
A;Molecule type: DNA	
A;Residues: 1-434 <NUX>	
A;Cross-references: EMBL:U55376; NID:91280130; PID:91280131; PIDN:AAA98003.1;	
A;Experimental source: strain Bristol N2; clone F16H11	
C;Genetics:	
A;Gene: CESP:F16H11.3	
A;Map position: X	
A;Introns: 49/3; 94/2; 190/1; 256/3; 326/3; 392/2	
C;Superfamily: Caenorhabditis elegans hypothetical protein ZK809.4	
Query Match	98.7%
Best Local Similarity	99.1%
Matches	430; Conservative 0; Mismatches 4; Indels 0; Gaps 0
Qy	1 MVITNRNSNTYAVEQAFAPRDKYNTVYLWLVLFVGVLPPWNMFITTAPEYYNNWFKPDG 60
Db	1 MVITNRNSNTYAVEQAFAPRDYNTVYLWLVLFVGVLPPWNMFITTAPEYYNNWFKPDG 60
Qy	61 VETWYSKEPMGLTISQLPNASINVNLFLIIAGLPIYRFAPVCNIVNLTLILYI 120
Db	61 VETWYSKEPMGLTISQLPNASINVNLFLIIAGLPIYRFAPVCNIVNLTLILYI 120
Qy	121 VLEPTEDSISWFWTYFLGMATSINFSNGLYENSVGFISLVLVCAATALFIFTKQDFYHYHHOKGMEIREKA 180
Db	121 VLEPTEDSISWFWTYFLGMATSINFSNGLYENSVGFISLVLVCAATALFIFTKQDFYHYHHOKGMEIREKA 180
Qy	181 TVVKIGTVYFLNDEPKLVAIVYFGISLVLVCAATALFIFTKQDFYHYHHOKGMEIREKA 240
Db	181 TVVKIGTVYFLNDEPKLVAIVYFGISLVLVCAATALFIFTKQDFYHYHHOKGMEIREKA 240
Qy	241 ETDRPSPLWLTTFTNCYGOLENWFCAVLTIFPVMMTYTRGDSFLNKIMSENDEI 300
Db	241 ETDRPSPLWLTTFTNCYGOLENWFCAVLTIFPVMMTYTRGDSFLNKIMSENDEI 300
Qy	301 YLTLSFLNFLFAAIGTIVASKIWLPTPRYLKFAITLRAFLPTFFFCNYRVQTRAPV 360
Db	301 YLTLSFLNFLFAAIGTIVASKIWLPTPRYLKFAITLRAFLPTFFFCNYRVQTRAPV 360
Qy	361 EEEFDTEVCGIAMSSESHGTYSALAMGYTPNVPSHYSRPAQOVSUTLWGLLPGCILW 420

R;Swinburne, J
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19753
A;Accession: T23528
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-450 <WIL>
A;Cross-references: EML:Z79601; PIDN:CAB01882.1; GSPDB:GN00028; CESP:K09A9.3
A;Experimental source: clone K09A9
C;Genetics:
A;Gene: CESP-K09A9.3
A;Map position: X
A;Introns: 59/3; 115/2; 147/1; 318/1; 346/3
C;Superfamily: Caenorhabditis elegans hypothetical protein ZK809.4

RESULT 2

T28006 hypothetical protein ZK809.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T28006
R;Dobson, R.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z20454
A;Accession: T28006
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-461 <WIL>
A;Cross-references: EML:Z66303; PIDN:CAA92642.1; GSPDB:GN00022; CESP:ZK809.4
C;Genetics:
A;Gene: CESP-ZK809.4
A;Map position: 4
A;Introns: 54/3; 110/2; 142/1; 357/3
C;Superfamily: Caenorhabditis elegans hypothetical protein ZK809.4

Query Match 30.8%; Score 706.5; DB 2; Length 450;
Best Local Similarity 36.0%; Pred. No. 2.7e-48;
Matches 157; Conservative 88; Mismatches 164; Indels 27; Gaps 8;

Qy 1.2 VEQBA-FPRDKINIVNVLVILVGFGVLLPWNMFITIAPEYYNNWF----- RPD 59
Db 21 VEETPEEDKGNLVVFYILLHGIGTLMWNMLITISDFESYKMLANSTIDMDTGKVT 80
Qy 60 GWTWYSKEFMGSUTIGSOLNPNASINVNLFLIAGLIYRFAPVCENIVNLTLILLY 119
Db 81 GDTVYSSNFQSFQTIASQVNPNLNLNUNIFVKGGLAGRI-TVGTSIVAVCVITMI 138
Qy 120 IVLEPTEDSMSNFFWWFVILGMAINSFNSNGLYEMSVYGVGGDPHTYIGALLIGNNCGLL 179
Db 139 FLYVETSWLWGFTELITLITLVNGANGVYQNSTFGLASELPFKTNAVIGNNUCCGF 198
Qy 160 ITVVKI--GTVYFLNDEPKLVAIVYEGISLVLVCAIALEFFITQDFYHHHQKMEI 236
Db 199 VTLISMSTKAVPNLILDR--SFAYSSIALTIVCEFSFHLLKORFYQYSTRAERQ 254
Qy 237 R--EKAETDRPSPSILWTTFTNCYQGFAYLTIFP-VMMVTTRGDSGFLNK 292
Db 255 RAKNEAADNEGKMANVATFREAPQQLINVLVFFVTLSPFPGVMMYVKDEKKGTYYDF 314
Qy 293 IMSENDETYTLLTSELYENLFAAIGSTVASKIHWPTPYLKFAILRLALFTPEFFFCNYR 352
Db 315 PFLQNI--YFMDVTFLOPNVAFIGSIVAGRQWPANKWLWVPLVRLYLTFEFICNYL 372
Qy 353 VOTRAYPPFFESTDIFTYIGIAMSFSHGYLSALAMGYTPNVVPSHYSRFAAQLSVCTLMV 412
Db 373 PETRTLPIVFEESTWLFLITVAAMSFGSGYFSGLSMAYTSKTVDPSKAQVAGMMAGFFLIS 432
Qy 413 GLITGGIIPPVVIEHFV 428
Db 433 GIVSGLIFTMVIKFVV 448

RESULT 4

T2247 hypothetical protein K02E11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Accession: T23247
R;McMurray, A.
A;Submitted to the EMBL Data Library, July 1996
A;Reference number: 219715
A;Accession: T23247
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-800 <WIL>
A;Cross-references: EML:Z77665; PIDN:CAB01223.1; GSPDB:GN00023; CESP:K02E11.1
A;Experimental source: clone K02E11
C;Genetics:
A;Gene: CESP-K02E11.1
A;Map position: 5
A;Introns: 63/3; 80/3; 120/3; 137/1; 178/1; 280/3; 333/3; 418/3; 464/2; 496/1; 521/2;
C;Accession: T23528
hypothetical protein K09A9.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C;Accession: T23528

RESULT 3

hypothetical protein K09A9.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C;Accession: T23528

Db	273	FP-----GPIAENLKKSQLQSMYPIILLITVYNISDFVGKSLTALYLWQSIKSAT	321	A; Reference number: JC4195; MUID: 95367016 A; Accession: JC4195 A; Molecule type: mRNA A; Residues: 1-327 <WIL> C; Cross references: EMBL:X86682; NID:9951302; PIDN:CAA60381.1; PID:9951303 C; Comment: This protein has a role in the growth response and participates in the pro C; Genetics:
Oy	334	FAILRLALFIPFFFCNYRVTQRAYPVFFESTDIFVIGGIAMMSFSHGYLSALAMGYTPNV	393	A; Gene: der12 A; Start codon: AUG C; Keywords: nucleoprotein; phosphoprotein C; Cross references: EMBL:X86682; NID:9951302; PIDN:CAA60381.1; PID:9951303 F; 80-157:Domain: hydrophilic #status predicted <HYD> F; 54/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted F; 107/Binding site: phosphate (Thr) (covalent) (by casein kinase II) #status predicted
Db	322	WACIVRLFYPLESAC-----LRGPWKRLRTIEVPPVVLTLEMGLTNGYLTSVLMIMAPKT	375	Query Match 9.1%; Score 209.5 ; DB 2; Length 327; Best Local Similarity 22.4% ; Pred. No. 2.0e-19 Matches 76; Conservative 49; Mismatches 134; Indels 81; Gaps 12;
Qy	394	VPSHYSRFAAQQLSVCSTMVGLLTTG----LW	420	Qy 131 WFFWVTLGMATSINFNSNLVENSVYGVGGDFPHTYIGALLIGNNCIGLITVVKI---- 185 Db 5 WF-----INSFCAVLQSSLFGQIGTMSTYSTFLSGOGLAGIAFAALAMLMLAS 54
Db	376	VHSEAEEAAIFPNMVFELGLGCGSVICWLW	406	Db 109 AELOADKNGPKPISPOQASPTLDLPEKEPBPPEPQPKBSVFWVERKINITALCLVL 168
RESULT 7	JC4196	36k hydrophobic nucleolar protein - human		Qy 186 GYTYFLNDEPKLVAYVF--GHSVLVLYCATL-----FFI-----TK 222 Db 55 GV-----DAQTSALGYFIFTPCVIGLISIVCYLSLPHLKARYLVTEKLSQAPTQELETK 108
C; Species: Homo sapiens (man)	C; Date: 10-Sep-1995 #sequence_revision 27-Oct-1995 #text_change 21-Jul-2000		Qy 223 QDFYHYHHQKGMEIR-----EKAETRSPSPSPLWTFTCNYGQLFNFVNWF 266 Db 109 AELOADKNGPKPISPOQASPTLDLPEKEPBPPEPQPKBSVFWVERKINITALCLVL 168	
R; Williams, J.B.; Lanahan, A.A.	R; Williams, J.B.; Lanahan, A.A.		Qy 267 CFAVTLTLPFPM-MTVTRGDSG---FLNKIMSENDEIYTLLTSLVFNLFAAIGSIV 320 Db 169 VFTVTSYFPAALTAMVTTSSNPGKNGLFENFI-----CCPFLFNYMDWLGRLSL 217	
Biochem. Biophys. Res. Commun. 213, 325-333, 1995			Qy 321 ASKIHW-----TRYLKPAILRALFIPFFENRYROTAYPVFFFESTDIFVIGGIAMSF 377 Db 218 TSYFLWPEDSQEDSQQLPLVCLRFLEYFLMCLHVQHAR-LPITFRODAYFTEMLLFAV 276	
A; Title: A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucle	A; Reference number: JC4195; MUID: 95367016		Qy 378 SICGYLSLAMGTAMGTVYSPHRYAAQSVCYTMVGLITG 417 Db 277 SNGYLVSLTMCLAPRQLPHEREVAGALMTFFLAGLSCG 316	
A; Accession: JC4196	A; Accession: JC4196		RESULT 9	
A; Molecule type: mRNA			hypothetical protein F36H2.2 - Caenorhabditis elegans	
A; Residues: 1-326 <WIL>			C; Species: Caenorhabditis elegans C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999	
A; Cross references: EMBL:X86681; NID:9951266; PIDN:CAA60380.1; PID:9951267			C; Accession: T21887 C; Accession: R; Steward, C. C; Submitted to the EMBL Data Library, October 1996 A; Reference number: Z19483 A; Accession: T21887 A; Status: preliminary; translated from GB/EMBL/DDBJ	
A; Experimental source: heart			A; Molecule type: DNA A; Residues: 1-513 <WIL> A; Cross references: ENBZ:Z1078; PIDN:CA03075.1; GSPDB:CN00019; CESP:F36H2.2 C; Genetics: A; Gene: CESP:F36H2.2 A; Map position: 1 A; Introns: 28/3; 56/1; 93/2; 163/3; 277/3; 314/1; 341/3; 384/2; 407/2; 458/2; 480/3	
C; Comment: This protein has a role in the growth response and participates in the prolif	C; Comment: This protein has a role in the growth response and participates in the prolif			
C; Genetics:	C; Genetics:			
A; Gene: der12	A; Gene: der12			
A; Start codon: AUG	A; Start codon: AUG			
C; Keywords: nucleoprotein; phosphoprotein	C; Keywords: nucleoprotein; phosphoprotein			
F; 54/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted	F; 107/Binding site: phosphate (Thr) (covalent) (by casein kinase II) #status predicted			
A; Cross references: EMBL:X86681; NID:9951266; PIDN:CAA60380.1; PID:9951267	A; Cross references: EMBL:X86681; NID:9951266; PIDN:CAA60380.1; PID:9951267			
A; Title: A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucle	A; Title: A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucle			
Db	143	INFNSNLVENSYGVGGDPAPHYIGALLIGNCGLLITVVKI-----GYTYFLNDPPLK 197		Qy 278 MMVTTTREDGSFLINKINSEDTYLTSLFNLFAIGSVASKTHW-----TPRYLKFA 335 Db 181 ITAMVTSST-----PGKWSOFFNPICCPFLNMDWLGRLLTSFLWPDDEDSRLPLL 234
Qy	198	VALVF---GISLVLVLYCATAIALLFFITKQDFY---HYHHQKGMEIRKA----- 240		Db 121 SSPQKVALTDLDEKEPESEDEPQPKPVSFTVQPKINLTALCLVLFVTVLSVFP 180
Db	61	SALGYFTPYVGILMSVCLYSLPHKFARYLANKSSQAQQLETKAELLQSDENGIP 120		Qy 278 MMVTTTREDGSFLINKINSEDTYLTSLFNLFAIGSVASKTHW-----TPRYLKFA 335 Db 181 ITAMVTSST-----PGKWSOFFNPICCPFLNMDWLGRLLTSFLWPDDEDSRLPLL 234
Qy	241	-----ETDRPSPSPLWTFTCNYCQLFNNWFCAFVFTLTFPV 277		Db 181 ITAMVTSST-----PGKWSOFFNPICCPFLNMDWLGRLLTSFLWPDDEDSRLPLL 234
Db	121	SSPKVQVALTDLDEKEPESEDEPQPKPVSFTVQPKINLTALCLVLFVTVLSVFP 180		Qy 336 IILRALFIPFFFCNYRVTQRAYPVFFESTDIFVIGGIAMSSHGYLSALAMGYTPNVVP 395
Qy	278	MHMVTTTREDGSFLINKINSEDTYLTSLFNLFAIGSVASKTHW-----TPRYLKFA 335		Db 235 VCFRLFVPLMICVQFQSR-LPILFPLQDAYFTIFMFLFAVNSGVLVLSLTMCAPRQLV 293
Db	181	ITAMVTSST-----PGKWSOFFNPICCPFLNMDWLGRLLTSFLWPDDEDSRLPLL 234		Db 294 PHREVAGALMTEFLALGLSCG 315
Qy	336	IILRALFIPFFFCNYRVTQRAYPVFFESTDIFVIGGIAMSSHGYLSALAMGYTPNVVP 395		RESULT 8
Db	235	SHYSRFAAQQLSVCSTMVGLLTTG 417		JC4195 36k hydrophobic nucleolar protein - mouse
Qy	396	VCFRLFVPLMICVQFQSR-LPILFPLQDAYFTIFMFLFAVNSGVLVLSLTMCAPRQLV 293		C; Species: Mus musculus (house mouse) C; Date: 03-Sep-1995 #sequence_revision 27-Oct-1995 #text_change 05-Nov-1999
Db	294	PHREVAGALMTEFLALGLSCG 315		C; Accession: JC4195 R; Williams, J.B.; Lanahan, A.A.
RESULT 8	JC4195			
36k hydrophobic nucleolar protein - mouse	36k hydrophobic nucleolar protein - mouse			
C; Species: Mus musculus (house mouse)	C; Species: Mus musculus (house mouse)			
C; Date: 03-Sep-1995 #sequence_revision 27-Oct-1995 #text_change 05-Nov-1999	C; Date: 03-Sep-1995 #sequence_revision 27-Oct-1995 #text_change 05-Nov-1999			
A; Title: A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucle	A; Title: A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucle			

Qy 81 NASINVNLILLIAGPL-----IYRVEAPVCFNIVNLII----LILVIVLEPT----- 125
 Db 62 TELOANFEATVYVSIACVGLSILNVPAK--SLSNSPRMIVGILVVVFPIITALF 119
 Qy 126 --EDSMSWPFWTFLGMSLTINF-SNGLYENSVYVGDDPHTVIGALLIGNNICGLLIT 181
 Db 120 WNFDDQTFFFLNLSMLISIACFGSILGMAGGVGLSALFP SOYTQAVWYQSEFAGVLA 179
 Qy 182 VVKI --GVT--YFLNDEPKLVATYFGISIIVLVAIALEFFT-----KODE 225
 Db 180 LMSILCQAVTSVDLNGQ----MFPGFSLIMCPISLATYYYLTLTPPMITDDSEG 233
 Qy 226 YTHHOKGMR-----EKAENDRPSPSLWTTFTNCYGOLFNWWFCFAVTL 272
 Db 234 IENEVEEVSTIAQANHFPPTIDSNSCQEEHOLPK ---WTMYDILRK-----SAIDL 282
 Qy 273 TIFPYMMTVTTRGDSGFLNKIMSEN---DEIYTLLTSLVFLNFAAIGTISVASKIHW 327
 Db 283 TTISVLVLTIAAYGTLTSVHSTSRRNHTWNSYFSAVASELYNVGDGRSSANSRL- 341
 Qy 328 TPRVLFKAIIIRLFIPPEFFCN-----YRVQTRAYPVFFESTD-----TFVIG 371
 Db 342 SPKYLIIISPLREALPIIAMCNVVEHILMHFLHMTRVFLFLSIDKDSRELAGSIIALF 401
 Qy 372 GIAMS-----FSHGYLISALAMGYT 390
 Db 402 GVTAAIFISFFCINFCISYYT 425

RESULT 10
 S336712 FUN26 protein - yeast (Saccharomyces cerevisiae)
 N;Alternative names: Protein YAL022C
 C;Species: Saccharomyces cerevisiae
 C;Accession: S36712
 R;Ouellette, F.; Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Zeng, B.; Fortin, N.; D
 Submitted to the EMBL Data Library, January 1993
 A;Description: Sequencing of chromosome I from Saccharomyces cerevisiae: analysis of a 3
 R;Reference number: S36711
 A;Accession: S36712
 A;Molecule type: DNA
 R;Residues 1-517 <QUE>
 A;Cross-references: EMBL:L05146; NID:g171851; PIDN: AAC04935..1; PID:g171853; MIPS: YAL022C
 C;Genetics

A;Gene: FUN26
 A;Cross-references: SGD:S0000020; MIPS: YAL022C
 A;Map position: 1L
 C;Keywords: transmembrane protein

F;79-193/Domain: transmembrane #status predicted <TM1>
 F;155-171/Domain: transmembrane #status predicted <TM2>
 F;218-234/Domain: transmembrane #status predicted <TM3>
 F;245-261/Domain: transmembrane #status predicted <TM4>
 F;344-360/Domain: transmembrane #status predicted <TM5>
 F;412-428/Domain: transmembrane #status predicted <TM6>
 F;445-461/Domain: transmembrane #status predicted <TM7>
 F;500-516/Domain: transmembrane #status predicted <TM9>

Query Match 8.0%; Score 182.5; DB 2; Length 517;
 Best Local Similarity 19.9%; Pred. No. 6.3e-07;
 Matches 79; Conservative 69; Mismatches 131; Indels 117; Gaps 20;

Qy 23 NIVYWLVLVFGVLLPWNMNETIAPEYYNNWFKEPDGVETW--YSKEFMSLTIGSQL 79
 Db 75 NLSYITTEFAIGLWPNCLTS-ASQYFKHDIFKDTSI--WAKIETSSNMMSFSTSSML 131
 Qy 80 PNASINVNLIL-----IAGPLTR--VAPAVCINVNLTILIVIVLEPTD 127
 Db 132 -----FNIYLAKROYKYSRRVING-LVNEETIVTVMCF---FVILHIL----- 171
 Qy 128 SMSMWPFWTFLGMSLTI-NFSNGLYENSVYGGDFPHTYICALLIGNNIGLLITVKG 186

Db 172 -PKWFNFEMFTMMLVYVISMGCTAMTONGIMAIANVFGSEYSGQAVMVQGAVVLPSSLVLF 230
 Qy 187 VTYFLNDEPKLVK--IVYFGISLVLVLCIALFFITK----- 222
 Db 231 LAFIENS SVSTGGTLLYFFFTTLLVWVTC-VVMFSVSKISRKVNVNWEDGHITDVLLG 289
 Qy 223 -----QDFYHYH-----HQKMEIREKAETDRPSPSLWTTFTNC 257
 Db 290 SLRSNEEIRIVGRIDQMEDEDHRTNTGTRDDNCEQLKVPFE-----VLFALK 341
 Qy 258 YGOLFVNWFVFCFAVTLTIFPYNMVTTRGDSGFLNKIMSENDEIYLTTSFLVFNLF 317
 Db 342 KYLVLSIFTFVVTL-VFPVFASAT-----YVTGLPLSNAQYIPLI--FTLWNLQDLYG 392
 RESULT 11
 T22164 hypothetical protein F44D12.9 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #text_change 15-Oct-1999
 R;Coles, L.
 submitted to the EMBL Data Library, December 1995
 A;Reference number: Z19525
 A;Accession: T22164
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Cross-references: EMBL:Z68298; PIDN:CAA92605..1; GSPDB:GN00022; CESP:F44D12.9
 A;Experimental source: clone F44D12
 C;Genetics:
 A;Gene: CESP:F44D12.9
 A;Map Position: 4
 A;Introns: 19/1; 37/2

Query Match 7.8%; Score 180; DB 2; Length 143;
 Best Local Similarity 40.0%; Pred. No. 2.4e-07;
 Matches 32; Conservative 18; Mismatches 26; Indels 4; Gaps 1;

Qy 14 QEAEPDKYNNVWYLYLGWPMFETTAPEYYNNWFKEPDGVETWYSKEFMSLT 73
 Db 67 KESSPPTRWLYIITPHRGMLMSWNMFITIAPOYHYDWFN---NTNYQDSEMSII 122
 Qy 74 TIGSQLPNASINVNLIL 93
 Db 123 GVTSQLPNVGMILNTVVM 142

RESULT 12
 D85064 hypothetical protein AT4905120 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: D85064
 R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
 A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A;Reference number: A85001; MUID: 20083488
 A;Accession: D85064
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-418 <STO>
 A;Cross-references: GB:NC_001268; PIDN:g7267271; PION: CAB81054..1; GSPDB: GN00140
 C;Genetics:
 A;Gene: At4g05120
 A;Map position: 4

Query Match 6.3% Score 144; DB 2; Length 418;	Qy 187 VTVYFL---NDEPKLVAIVYFGISLVLVCAIALFFI-TKQDFYHYHHQKG-----233
Best Local Similarity 20.6% Pred. No. 0.00054; Mismatches 159; Indels 134; Gaps 23;	Db 168 ITRAFAERSNNSLRKGMANIFLAISTFIELLCVLYAVFPKLPIVKYRRKAASEGSKTV 227
Matches 96; Conservative 76; Mismatches 159; Indels 134; Gaps 23;	Qy 234 -----MTEREKAETD-----RPSPLWTTFTNCYGOFLNWFCEFAVLTIFPVPM 278
Qy 18 PRDKYNIVYLVI--LYGFVLLPWNMFITIAPEYY-VNYWFKPDPGETWYSKEF-MGS 72	Db 228 VADLAAAGTQNLSLSDDSKNMLRKELLQN----IDHAVNLFLIVLTLSPF--280
Db 9 PPEKLGKIQAMVCCILSIGLSWSNMLTADYYKKVFPDHPSPRLTVYQPFAFGT 68	Qy 279 MTYTRGSGFLNKIMSEN-DEIYTLI-TSFLVNFIAAGTSIVASKIHW--PPRY 331
Qy 73 LTI---GSQLPNNSINVNLFLIAPIYRYFAPVCNFNIINLILVILEPIEDS 128	Db 281 -----GFLYENTGGHGLGWYALLVATYNEWDFGRYAPLV---KWLKLENRKA 327
Db 69 ILILAYHESKINTNRKRNLIYI-----FTSTFLVLD----103	Qy 332 LKPAIIIRALTFPPEFFFNYRVOTRFVFFFESTDILVIGGAMTSIGYTP 391
Qy 129 MSWFPEWVTLMAT-----SINFNSGLYENSYVG--VGG-----DEPHTYWIG 167	Db 328 LTAVLTLRFLYFLPAPFYTAKGDKGMIML/S-----ILGLTTGHLYCIMTAP 377
Db 104 -----LATKRGGGIPIGYIGLCAVVASFLADAVYQGGMIGDLSLMCPLVQSFMG 153	Qy 392 NVYPSHYSRFAAQLSVCTLM---VGLTGGLN 420
Qy 168 ALIGNNICCLITVVKIGTVYFLNDEPKLVAIVYFGISLVLVCAIALFFI-TKQDFY 226	Db 378 NGKGPKERNALGNLWVFILGGAVGVLGWLW 410
Db 154 GLAVSGALTSAHLRLLTAKAFK-TNDGPRKGAMMFLAISTCIELLCVLYAVFPKLPIV 212	
Qy 227 HYHQKGMEIREK-----AETRDRPSPLWTTFTNCYGOLENW 265	
Db 213 KYYRKAASEGSKTVSADLAAGIQNQSDDLSKNRSLSKRNLLIONIDAVNLFLY 272	RESULT 14 H6383 Na+ transporter - Methanococcus jannaschii C:Species: Methanococcus jannaschii C:Date: 13-Sep-1996 #sequence-revision 13-Sep-1996 #text-change 31-Dec-2000 C:Accession: H64383 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, erson, J.D.; Sadow, P.W.; Hanaa, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996 A:Authors: Kaine, B.P.; Bordovskiy, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannas A:Reference number: A64300; MUID:96333999 A:Accession: H64313 A:Status: preliminary; nucleic acid sequence not shown; translation not shown A:Molecule type: DNA A:Residues: 1-432 <BUL> A:Cross references: GB:U67314; GB:L77117; NID:92826304; PIDN:AAB98666.1; PID:91591386 C:Genetics: A:Map position: REV598993-597695 C:Superfamily: probable transporter MJ0672
Qy 266 FCFAVLTIFPVMMTVTRGDSFLNLKIMSEN-DEIYTLFLSFLVLENFAAIGTSIVASK 323	Query Match 5.8% Score 134; DB 2; Length 432; Best Local Similarity 19.1%; Pred. No. 0.0034; Matches 82; Conservative 80; Mismatches 133; Indels 134; Gaps 22;
Db 273 VC---TLSIIP-----GFLYENTGQHGDNAYALV-VAMYNCWDLVGR----312	Qy 66 SKEFMG-----SLTIGSOLPNASINVNFLIITAGPLIYRFAPCFNIVNLNTILILY 119 Db 4 SKFEGIQLITASLIFSSLPDLYKG---FWVTLG--MATSINSN-GLYENSVYGVGDFPHT----61
Qy 324 IHWPPTPRYLK-----FAITLRLALPFPPFFCNYVQTRAVPVEFESTDIFVIGGIAM 375	Db 62 GIFN-LKEALITYFAHPFLFLGGFLMAQALKNHLDKFIAKLLNYKDPTTTCFLMFL 120
Db 313 -YTPLVKWLKENRKLTIAVLRSYLLIPAPIFY-TAKYGDQGMIML----1SVLG----362	Qy 165 --YIGALIIGNNCIGLLTVVKIGTVYFLNDEPKLVAIVYFGISLVLVCAIALFF-----FLNDEPRL 197
Qy 376 SFSGHYSALAMGTYPNSVPSHYSRFAAQLSVCTLMYGLLGTGLW 420	Db 121 SAYFLSMWISNTSATLILPITALGLLHKTKGKLRFELLGAYASASIGGIATIGSPNA 180
Db 363 -LINGHLTVICM---TIAPKGY-KGPEQNALAQNLLVFLGGIF 401	Qy 198 VAIVY-----FITKQDFYHYHHQKGMEIREKA 240 Db 181 IASSYLDYGFSSWFKVGPISLLFLICLTLYIYFKWIKPED-----IAQARM 231
RESULT 13	Qy 241 ENDRPSPI-----LWTFNTNCYGOLENWWECAVLTIFPVMMTVTRGDSGFLN 291 Db 232 ELSRNAYKLLVFLIASW-TISDYLESEIENQY-FDSVIAIFIATL-----FVF 281
E885054 hypothetical protein At4g05130 [imported] - Arabidopsis thaliana	Qy 292 KIMSEND---EYVLL-----TSPLVFNFAAIGSTIVASKHMPPTP 330 Db 282 NLVEVNDFKKIDWGTLLFGLGVGDLSLMCPELIOQSMSGMAGALTsvrl- TPI 333
C:Species: Arabidopsis thaliana (mouse-ear cress)	
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001	
C:Accession: E85064	
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999	
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.	
A:Reference number: A85001; MUID:20083488	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-418 <STD>	
A:Cross-references: GB:NC_001268; NID:97267272; PIDN:CAB81055.1; GSPDB:GN00140	
C:Genetics:	
A:Gene: At4g05130	
A:Map position: 4	
Query Match 6.0% Score 138.5; DB 2; Length 418;	
Best Local Similarity 21.0% Pred. No. 0.0015; Mismatches 73; Indels 111; Gaps 23;	
Matches 95; Conservative 73; Mismatches 73; Indels 111; Gaps 23;	Qy 21 KYNEVYWLIVLGVFVLLPWNMFITIAPEYY-VNYWFKPDPGETW-YSKERNGSLTIGSQ 78
Db 16 KYQAMVVCIL-SIGSFWSNMLTADYYKKVFPDHPSPRLTYQPIALGTM----71	Qy 79 LPNASINVNLFLIAPIYRYFAPVCNFNIINLILVILEPTEDSMWFFWVTLG 138
Qy 72 LAYRESKISTRKRILTYGLF-----TISTFLVLDL-----TKG 108	Db 72 LAYRESKISTRKRILTYGLF-----TISTFLVLDL-----TKG 108
Qy 139 -----MATISINFNSLYENSYV-GVGGDFP---HTYIGALLIGNNICGLLITYVVKIG 186	Qy 292 KIMSEND---EYVLL-----TSPLVFNFAAIGSTIVASKHMPPTP 330
Db 109 HGGIGHIVLCTIVASFLADAVYQGLGVGDLSLMCPELIOQSMSGMAGALTsvrl- 167	Db 282 NLVEVNDFKKIDWGTLLFGLGVGDLSLMCPELIOQSMSGMAGALTsvrl- TPI 333

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: February 27, 2002, 16:56:28 ; Search time 14.4 Seconds
(without alignments)
678.224 Million cell updates/sec

Title: US-08-816-011f-63
Perfect score: 2294
Sequence: 1 MWIINRSNTYAVEESEAPRDF.....LTGGLWPVVIBHFVDKPSIL 434

scoring table: BLOSUM62
Gapp 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:
 1: /cgn2_6/picodata/2/iaa/5a_COMB.pep:
 2: /cgn2_6/picodata/2/iaa/5b_COMB.pep:
 3: /cgn2_6/picodata/2/iaa/6a_COMB.pep:
 4: /cgn2_6/picodata/2/iaa/6b_COMB.pep:
 5: /cgn2_6/picodata/2/iaa/PCTUS_CORG.pep:
 6: /cgn2_6/picodata/2/iaa/backfile1.pep:
 *

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description	
1	343.5	15.0	456	4	US-09-058-389A-4	Sequence 4, Appli	
2	326.5	14.2	456	4	US-09-058-389A-2	Sequence 2, Appli	
3	214	9.3	326	4	US-09-058-389A-3	Sequence 8, Appli	
4	181	7.9	247	4	US-09-058-389A-8	Sequence 4, Appli	
5	111.5	4.9	398	1	US-08-057-938-4	Sequence 4, Appli	
6	7	111.5	4.9	398	1	US-08-076-000-4	Sequence 4, Appli
7	111.5	4.9	398	1	US-08-072-840-4	Sequence 4, Appli	
8	111.5	4.9	398	2	US-08-076-976-4	Sequence 4, Appli	
9	111.5	4.9	398	3	US-08-074-410-4	Sequence 6, Appli	
10	111.5	4.9	398	4	US-08-086-673B-4	Sequence 4, Appli	
11	109.5	4.8	397	4	US-08-086-673B-6	Sequence 6, Appli	
12	108.5	4.7	398	1	US-08-07-938-6	Sequence 6, Appli	
13	108.5	4.7	398	1	US-08-076-000-6	Sequence 6, Appli	
14	108.5	4.7	398	1	US-08-072-840-6	Sequence 6, Appli	
15	108.5	4.7	398	2	US-08-076-976-6	Sequence 6, Appli	
16	108.5	4.7	398	3	US-08-074-410-6	Sequence 6, Appli	
17	108.5	4.7	398	4	US-08-086-673B-6	Sequence 6, Appli	
18	106.5	4.6	395	1	US-08-07-938-2	Sequence 5, Appli	
19	106.5	4.6	395	1	US-08-07-938-5	Sequence 5, Appli	
20	106.5	4.6	395	1	US-08-076-000-2	Sequence 2, Appli	
21	106.5	4.6	395	1	US-08-076-000-5	Sequence 5, Appli	
22	106.5	4.6	395	1	US-08-072-840-2	Sequence 2, Appli	
23	106.5	4.6	395	1	US-08-072-840-5	Sequence 5, Appli	
24	106.5	4.6	395	2	US-08-076-976-5	Sequence 2, Appli	
25	106.5	4.6	395	2	US-08-076-976-5	Sequence 5, Appli	
26	106.5	4.6	395	3	US-08-074-410-2	Sequence 2, Appli	
27	106.5	4.6	395	3	US-08-074-410-5	Sequence 5, Appli	

ALIGNMENTS

RESULT 1
US-09-058-389A-4
; Sequence 4, Application US/09058389A
; Patent No. 6130065
; GENERAL INFORMATION:
 : APPLICANT: Belt, Judith A.
 : BELT, Judith A.
 : ATTORNEY: Crawford, Charles R.
 : CRAWFORD, Charles R.
 : TITLE OF INVENTION: A NITROBENZYL MERCAPTOPURINERIBOSIDE
 : TRANSPORT
 : (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
 : PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
 : USE
 : NUMBER OF SEQUENCES: 22
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: David A. Jackson, Esq.
 : STREET: 411 Hackensack Ave, Continental Plaza, 4th
 : STREET: Floor
 : CITY: Hackensack
 : STATE: New Jersey
 : COUNTRY: USA
 : ZIP: 07601
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC Compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentNet Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/058-389A
 : FILING DATE: April 9, 1998
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Jackson Esq., David A.
 : REGISTRATION NUMBER: 26,774Z
 : REFERENCE/DOCKET NUMBER: 1340-1-013N
 : TELEPHONE: 201-343-1684
 : TELEFAX: 201-343-1684
 : INFORMATION FOR SEQ ID NO: 4:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 456 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : DESCRIPTION: hENT1
 : HYPOTHETICAL: No
 : FRAGMENT TYPE:
 : ORIGINAL SOURCE:
 : ORGANISM: Homo sapiens
 : US-09-058-389A-4

Query Match 15.0% Score 343.5; DB 4; Length 456;
 Best Local Similarity 24.3%; Pred. No. 3.3e-26;
 Matches 112; Conservative 77; Missmatches 212; Indels 59; Gaps 13;

Match 18 PRDKYNIVWVLVLYFGFVLLPWNMFTTAIPEYYVNVNWFKPDGWE--TWSKEFMGSUT 74
 Db 7 PQDRKAVWLFFMLGTLPLWNMFATQYTNRLDMSQNTSVLTAELSDAQAAA 65

Query 75 IGSQLP- -NASINVNFLILLAGLIVRYFA- - - - -PYCFNIVN---LTILLI 117
 Db 66 PAAPLEPNRSLSAFNNYNTLCAMPLPLIFTYLNSFLHORIPQSVRILGSVAILLVELI 125

Query 118 LIVILEPTEDMSMSRFWWTLGMAITSINFNSNGLYENSYGVGGDFPHTYIGALLIGNNCIG 177
 Db 126 TAILVKVQLDALP FFWITMIKIVLINSFGAIIQLGSLSFLAGLPLASTAPIMSGQGLAG 184

Query 178 LTIVVKVIGTVYFLNDEPKLVATAVYFGISLVLIVCAIALEFLITKQDFHYHQ----- 231
 Db 185 FEASVAMI-CAIASSELSEASAFYTFITACAVILTTICLGLPRLEYRYCQOKLKGEP 243

Query 232 - - - - -KGMIREKAETDRPSPLWTT - - - - -FTNCYGLCPNFWFCFAVT 271
 Db 244 GEQETKLDLISKGEPRAGKEESVSVNSQSPINESHTIKALKNISLAFLSVCIFIT 303

Query 272 LTIFPVMMTWTTRGDSGFNLINKIMSENDEITYLTTSFLYENFLAAIGSIVASKIHWP--TP 329
 Db 304 IGMPAVTIVKVEKSSTAG---SSTWERYFPIVSCFLTNIDMLGRSLTAVMMPGKDS 358

Query 330 RYLFIAITLRLAFIPLVLLARLVPILLICN--IKPRRLTLYVEHDWAEFMAFAFSNGYTLASLCMC 388
 Db 359 RWLPLSLVLLARLVPILLICN--IKPRRLTLYVEHDWAEFMAFAFSNGYTLASLCMC 416

Query 389 YTPNVVPHYSRFAAQLSVCTLYMGLTGGLWPVIEHFV 428
 Db 417 FGPKVKPAEAETGAIMAFFCLGLGAVFSEFLFRATV 456

RESULT 2
 US-09-389A-2
 Sequence 2, Application US/09058389A
 GENERAL INFORMATION:
 APPLICANT: Patel, Divyen
 APPLICANT: Crawford, Charles R.
 APPLICANT: Bell, Judith A.
 APPLICANT: Crawford, Charles R.
 TITLE OF INVENTION: A NITROBENZYL MERCAPOURINERIBOSIDE
 TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
 TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
 TITLE OF INVENTION: USE
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David A. Jackson, Esq.
 STREET: 411 Hackensack Ave, Continental Plaza, 4th
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/058,389A
 FILING DATE: April 9, 1998
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson, Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 1340-1-013N

RESULT 3
 US-09-389A-3
 Sequence 3, Application US/09058389A
 GENERAL INFORMATION:
 APPLICANT: Bell, Judith A.
 APPLICANT: Crawford, Charles R.
 APPLICANT: Patel, Divyen
 TITLE OF INVENTION: A NITROBENZYL MERCAPOURINERIBOSIDE
 TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
 APPLICATION NUMBER: US/09/058,389A
 FILING DATE: April 9, 1998
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson, Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 1340-1-013N

STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/058, 389A
 FILING DATE: April 9, 1998
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 1340-1-013N
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 326 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: Protein hNRP36
 HYPOTHETICAL: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-09-058-389A-3

Query Match 9.3%; Score 214; DB 4; Length 326;
 Best Local Similarity 22.0%; Pred. No. 1.8e-13;
 Matches 71; Conservative 52; Mismatches 139; Indels 60; Gaps 8;

Qy 143 INFNSGLYENSYVGCGDFPHITYIGALLIGNNCGLITYVVKI----GYTYFLNDEPKL 197
 Db 7 INFSFAVLQGSLFGQLGTMPSYSTFLSGGLAGIFAAALMLMSMASGV----DAET 60
 Qy 198 VAIVYF---GISLVILVCAIALFFTKDQFY---HYHHHQKGMEIREKA-----240
 Db 61 SALGYFATPYGILMSIVCPLPHKAFARYLANKSSQAAQELETKAELQLSDENGIP 120
 Qy 241 -----ETDRPSPSILWTTFTNYGOLFNWWFCAVTLTIFPV 277
 Db 121 SSPQKVVALTLQDLEKEPESSPDPDEPKQPKPSVFTVQKTIWLTACLVLYETVTLSVFP 180
 Qy 278 NMVVTTRGDSEFLNKIMSENDEIYTULTSFLWEDDSRLPLL 235
 Db 181 ITAMVTSSTS----PGK4SOFPN1CCFLLENIMDWLGRSLTSYFLWEDDSRLPLL 234
 Qy 336 TILRALFTPFEEFCNYVRQTAAPYFESTIDIVIGCIAAMSQYTPNPNVP 395
 Db 235 VCLRFLEVPLMCHVQPRSR-LPILFQDANYTFLMELPAVSNGYLVLSDLMCLAPRQLV 293
 Qy 396 SHYSRFAAQSLSVCTLMVGLLTG 417
 Db 294 PHEREVAGALMTFFLAGLSCG 315

RESULT 4
 US-09-058-389A-8
 ; Sequence 8, Application US/09058389A
 ; Patent No. 6130065
 ; GENERAL INFORMATION:
 ; APPLICANT: Crawford, Charles R.
 ; APPLICANT: Patel, Divyen
 ; TITLE OF INVENTION: A NITROBENZYLMMRCAPTOPURINERIBOSIDE
 ; TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT

RESULT 5
 US-08-097-938-4
 ; Sequence 4, Application US/08097938
 ; Patent No. 5620174
 ; GENERAL INFORMATION:
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
 ; TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
 ; NUMBER OF SEQUENCES: 59

; TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
 ; TITLE OF INVENTION: USE
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: *IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/058, 389A
 FILING DATE: April 9, 1998
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 1340-1-013N
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 247 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 US-09-058-389A-8

Query Match 7.9%; Score 181; DB 4; Length 247;
 Best Local Similarity 23.6%; Pred. No. 2.5e-10;
 Matches 59; Conservative 42; Mismatches 107; Indels 42; Gaps 7;

Qy 18 PRDKYNVYVWLVLYVGFVLLPWNNMFTIAPEYY-----
 Db 7 PRDSYHVGISFFLGLQFLPWNNFTTAITYFQARLAGAGNSTARLSTNHTGPDAF- 65
 Qy 63 TWYSKEPMGSLTIGSOLPNASTINVNFLFLTAGPLIYRVFAPVCNFVNLTIIILIVL 122
 Db 66 -----NNNNWTVLLSQPLFLTLNSFLYQCVPENRILG---SLAIIILFLATAAL 116
 Qy 123 EPTEDSMWSWFWVNTLGMATSINFNSGLYENSYVGCGDFPHITYIGALIGNNICGLITYV 182
 Db 117 VKVDMSPGPFFSITMASVCFINSFSAWLQSGLFGQLGTMSTYSLFLSGQGLAGIFAAAL 176
 Qy 183 VKI----GYTYFLNDEPKLFTKQDFY---GISLVLVCAIALFFTKDQFY---HYHHQ 231
 Db 177 AMILSMSAVG-----DAETSALGYYFTPCVGILMSIVCYLSLPHKAFARYLANKSSQA 230

Qy 232 KGMEIREKAE 241
 Db 231 QAQELTKAE 240

CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 CITY: Washington, D.C.
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/097, 938
 FILING DATE: 26-JUL-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29, 959
 REFERENCE/DOCKET NUMBER: 22803-20006.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 398 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-097-938-4

Query Match 4.98; Score 111.5; DB 1; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041;
 Matches 75; Conservative 54; Mismatches 131; Indels 145; Gaps 18;
 US-08-476-000-4

Query 61 VETWYS-KEFMGSLTIGSQLPNAISINVNLFLITIAGPLIYRYPAVCNIVNLTLILV 119
 Db 56 VETEVSVDERSASVLTGK-----F 87
 Query 120 IYLEPTEDSMSWFF-----
 Db 88 VVGILPSNGMALWFLERTKKHPAVIYMANLALADLISVTFPLKIAHHNNWIYGEA 147
 Query 154 VYGVGGDFPHTYIGALLIGNICCOLITYVKGIVTYFLNDEP---KLVAVIYFGISLVI 209
 Db 148 LCNV-----LIG-FFGNYMCISLFWTCVQRYWIVNPNGHSRKKAIGISLAI 199
 Query 210 -LLVCAIAL-FITFKQDFYHYHHQKGMEIREKAETDRPSLWTFTNCY-----
 Db 200 WLLLVLVTPLYVVKQTIF----IPAL---NITTCHDVLPEQLLV 237
 Query 259 GOLFNVWFCFAVLTITFPPVMMTYYTRGDSGF-----NKIMSENDE-----IYT 302
 Db 238 GDMEFYFLSLAIGVFLPAPFLA-----SAYVLMIRLSSAMDENSESKKRRAIKLIVT 292
 Query 303 LITSFVLNFIAAIGSIYASKIHWPTPRLFIPFFFCCNYRVQTRAYPVFF 362
 Db 293 VLMYLI-CF-----TPSNL-----LLVWYHFLKSQGSSHVIALYI 328
 Query 363 ESTDIFVGGIAMSFSHGVL SALAMGTTPNVVPSPHYSREAAQLSV 407
 Db 329 VALCLSTLNCSIDBPVYFFVSHDFRDHAKNALLCRSVRTVKMQV 373

RESULT 6
 US-08-476-000-4
 Sequence 4, Application US/08476000
 Patent No. 5716789
 GENERAL INFORMATION:
 APPLICANT: SUNDELIN, JOHAN
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

RESULT 7
 US-08-472-840-4

TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 NUMBER OF SEQUENCES: 63
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZTP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,000
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/390,301
 FILING DATE: 25-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: ADLER, REID G.
 REGISTRATION NUMBER: 30,988
 REFERENCE/DOCKET NUMBER: 2803-0006-20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 398 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-476-000-4

Query Match 4.98; Score 111.5; DB 1; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041;
 Matches 75; Conservative 54; Mismatches 131; Indels 145; Gaps 18;
 Query 61 VETWYS-KEFMGSLTIGSQLPNAISINVNLFLITIAGPLIYRYPAVCNIVNLTLILV 119
 Db 56 VEVFSVDBFSASVLTGK-----
 Query 120 IYLEPTEDSMSWFF-----
 Db 88 VVGLPSNGMALWFLERTKKHPAVIYMANLALADLISVTFPLKIAHHNNWIYGEA 147
 Db 56 VEVFSVDBFSASVLTGK-----
 Query 154 VVGVGGDFPHTYIGALLIGNICCOLITYVKGIVTYFLNDEP---KLVAVIYFGISLVI 209
 Db 148 LCNV-----LIG-FFGNYMCISLFWTCVQRYWIVNPNGHSRKKAIGISLAI 199
 Query 210 -LLVCAIAL-FITFKQDFYHYHHQKGMEIREKAETDRPSLWTFTNCY-----
 Db 200 WLLLVLVTPLYVVKQTIF----IPAL---NITTCHDVLPEQLLV 237
 Query 259 GOLFNVWFCFAVLTITFPPVMMTYYTRGDSGF-----NKIMSENDE-----IYT 302
 Db 238 GDMEFYFLSLAIGVFLPAPFLA-----SAYVLMIRLSSAMDENSESKKRRAIKLIVT 292
 Query 303 LLTSFLVNLFAAIGSIYASKIHWPTPRLFIPFFFCCNYRVQTRAYPVFF 362
 Db 293 VLMYLI-CF-----TPSNL-----LLVWYHFLKSQGSSHVIALYI 328
 Query 363 ESTDIFVGGIAMSFSHGVL SALAMGTTPNVVPSPHYSREAAQLSV 407
 Db 329 VALCLSTLNCSIDBPVYFFVSHDFRDHAKNALLCRSVRTVKMQV 373

RESULT 6
 Sequence 4, Application US/08476000
 Patent No. 5716789
 GENERAL INFORMATION:
 APPLICANT: SUNDELIN, JOHAN
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

RESULT 7
 US-08-472-840-4

Query 303 LLTSLFVNLFIAIGSIVASKIHWPTPRYKFAILRLAFLIPFFCNRYQTRAYPVEF 362
 Database 293 VLAMYLI -CF-----TPSNL-----LLVHYFLIKSQGSQHVVAYLI 328
 Query 363 ESTDIFVIGGIAMSFSHGYSALAMGTYTPNVPSPHSYRFAOLSV 407
 Database 329 VALCLSTLNCSIDPFEVYFVSHDFRDHAKNALCRSVRTVKOMQV 373

RESULT 9
 US-08-474-410-4
 Sequence 4, Application US/08474410
 ; GENERAL INFORMATION:
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FORESTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,410
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REED G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006-20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-474-410-4

Query Match 4.9%; Score 111.5; DB 3; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 200 WLILLVTPPLVVKQTF-----IPAL---NITCHDVLPOLVV 237
 Qy 259 GOLFNYWFCFAVLTIFPVMMTVTRDSGFL-----NKIMSENDE-----LYT 302
 Db 238 GDNFNYFELSLAIGVFLPAELTA-----SAYLMIRMSAMDENSEKKRAIKLV 292
 Qy 303 LLTSLFVNLFIAIGSIVASKIHWPTPRYKFAILRLAFLIPFFCNRYQTRAYPVEF 362
 Db 293 VLAMYLI -CF-----TPSNL-----LLVHYFLIKSQGSQHVVAYLI 328
 Qy 363 ESTDIFVIGGIAMSFSHGYSALAMGTYTPNVPSPHSYRFAOLSV 407
 Database 329 VALCLSTLNCSIDPFEVYFVSHDFRDHAKNALCRSVRTVKOMQV 373

Query Match 4.9%; Score 111.5; DB 4; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 56 VETWYS-KEFMGLTIGSQLPNASINVNLFLIAGPLIYFAPVCFNIVNTLILV 119
 Qy 61 VETWYS-KEFMGLTIGSQLPNASINVNLFLIAGPLIYFAPVCFNIVNTLILV 119
 Db 56 VETVFSVDEFSASVLTK-----LTVFLPIVTIV-----F 87
 Qy 120 IYLEPTEDSMSWFF-----WVTLGMATSINSNGLYENS 153
 Db 88 VVGLPSNMAVLNVFLERIKKKHPAVIMANLAIDLSSVWPLKIAHYHNNWYGEA 147
 Qy 154 VYVGDDPFHTYIGALLGNICGLLITVKGTYFLINDEP-----KLVAIVYFGISLVI 209
 Database 148 LCNV-----LIG-FFYGNMCSILFMTCLSQRVWIVNMGHSHRKANIAIGISLAI 199

Query Match 4.9%; Score 111.5; DB 3; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 200 WLILLVTPPLVVKQTF-----IPAL---NITCHDVLPOLVV 237
 Qy 259 GOLFNYWFCFAVLTIFPVMMTVTRDSGFL-----NKIMSENDE-----LYT 302
 Db 238 GDNFNYFELSLAIGVFLPAELTA-----SAYLMIRMSAMDENSEKKRAIKLV 292
 Qy 303 LLTSLFVNLFIAIGSIVASKIHWPTPRYKFAILRLAFLIPFFCNRYQTRAYPVEF 362
 Db 293 VLAMYLI -CF-----TPSNL-----LLVHYFLIKSQGSQHVVAYLI 328
 Qy 363 ESTDIFVIGGIAMSFSHGYSALAMGTYTPNVPSPHSYRFAOLSV 407
 Database 329 VALCLSTLNCSIDPFEVYFVSHDFRDHAKNALCRSVRTVKOMQV 373

Query Match 4.9%; Score 111.5; DB 4; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 56 VETWYS-KEFMGLTIGSQLPNASINVNLFLIAGPLIYFAPVCFNIVNTLILV 119
 Qy 61 VETWYS-KEFMGLTIGSQLPNASINVNLFLIAGPLIYFAPVCFNIVNTLILV 119
 Db 56 VETVFSVDEFSASVLTK-----LTVFLPIVTIV-----F 87
 Qy 120 IYLEPTEDSMSWFF-----WVTLGMATSINSNGLYENS 153
 Db 88 VVGLPSNMAVLNVFLERIKKKHPAVIMANLAIDLSSVWPLKIAHYHNNWYGEA 147
 Qy 154 VYVGDDPFHTYIGALLGNICGLLITVKGTYFLINDEP-----KLVAIVYFGISLVI 209
 Database 148 LCNV-----LIG-FFYGNMCSILFMTCLSQRVWIVNMGHSHRKANIAIGISLAI 199

Query Match 4.9%; Score 111.5; DB 3; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 200 WLILLVTPPLVVKQTF-----IPAL---NITCHDVLPOLVV 237
 Qy 259 GOLFNYWFCFAVLTIFPVMMTVTRDSGFL-----NKIMSENDE-----LYT 302
 Db 238 GDNFNYFELSLAIGVFLPAELTA-----SAYLMIRMSAMDENSEKKRAIKLV 292
 Qy 303 LLTSLFVNLFIAIGSIVASKIHWPTPRYKFAILRLAFLIPFFCNRYQTRAYPVEF 362
 Db 293 VLAMYLI -CF-----TPSNL-----LLVHYFLIKSQGSQHVVAYLI 328
 Qy 363 ESTDIFVIGGIAMSFSHGYSALAMGTYTPNVPSPHSYRFAOLSV 407
 Database 329 VALCLSTLNCSIDPFEVYFVSHDFRDHAKNALCRSVRTVKOMQV 373

Query Match 4.9%; Score 111.5; DB 4; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 56 VETWYS-KEFMGLTIGSQLPNASINVNLFLIAGPLIYFAPVCFNIVNTLILV 119
 Qy 61 VETWYS-KEFMGLTIGSQLPNASINVNLFLIAGPLIYFAPVCFNIVNTLILV 119
 Db 56 VETVFSVDEFSASVLTK-----LTVFLPIVTIV-----F 87
 Qy 120 IYLEPTEDSMSWFF-----WVTLGMATSINSNGLYENS 153
 Db 88 VVGLPSNMAVLNVFLERIKKKHPAVIMANLAIDLSSVWPLKIAHYHNNWYGEA 147
 Qy 154 VYVGDDPFHTYIGALLGNICGLLITVKGTYFLINDEP-----KLVAIVYFGISLVI 209
 Database 148 LCNV-----LIG-FFYGNMCSILFMTCLSQRVWIVNMGHSHRKANIAIGISLAI 199

Query Match 4.9%; Score 111.5; DB 3; Length 398;
 Best Local Similarity 18.5%; Pred. No. 0.0041; Gaps 18;
 Matches 54; Mismatches 131; Indels 145; Gaps 18;

Db 200 WLILLVTPPLVVKQTF-----IPAL---NITCHDVLPOLVV 237
 Qy 259 GOLFNYWFCFAVLTIFPVMMTVTRDSGFL-----NKIMSENDE-----LYT 302
 Db 238 GDNFNYFELSLAIGVFLPAELTA-----SAYLMIRMSAMDENSEKKRAIKLV 292
 Qy 303 LLTSLFVNLFIAIGSIVASKIHWPTPRYKFAILRLAFLIPFFCNRYQTRAYPVEF 362
 Db 293 VLAMYLI -CF-----TPSNL-----LLVHYFLIKSQGSQHVVAYLI 328
 Qy 363 ESTDIFVIGGIAMSFSHGYSALAMGTYTPNVPSPHSYRFAOLSV 407
 Database 329 VALCLSTLNCSIDPFEVYFVSHDFRDHAKNALCRSVRTVKOMQV 373

STATE: D.C.
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIA TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006-20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-476-000-6

Query Match 4.7% Score 108.5; DB 1; Length 398;
Best Local Similarity 19.2%; Pred. No. 0_0083;
Matches 75; Conservative 59; Mismatches 141; Indels 115; Gaps 18;
US-08-472-840-6

QY 61 VETWYS-KEFMGSLTIGSQ---LPNASINVNPNUFLITAGPLIYRVA-----PVCFN 108
Db 56 VEIVFSVDEFSASVLTGKTTVFLPIVIVFVGPSNGMAIW-VELFRTRKKHPAVIY 114
QY 109 IVNUTILILVIVLEPTEDSMSHFWVTLGMATSINFNSGLYENSYVGDDFPHTYIGA 168
Db 115 MANLALADLSVT-----WFLKIHYHGNWNIGEALCNV-----LIG- 154
QY 169 LLTGNNICGLLTVVKIGTYFLNDEP---KLVATVFGISLVI-LLVCAAL-FFITK 222
Db 155 FFIGNMYCISLFLMTCLSVORYWIVNPMGHSKKANIAIGISLAIWLLVITPLYVK 214
QY 223 QDFYHYHHQKGMEIREKAETDRPSPSLWTFLTFINCY-----GOLFNYWFCFAVLT 273
Db 215 QTIF-----IPAL---NITTCHDVLPEOVLVGDMENYFLSLAIGVF 252
QY 274 IEPVMMTYTRGDSGF-----NKIMSENDE-----IYTLLTSFLVNLFAAIG 317
Db 253 LFP AFLTA-----SAYVLMIRMRSSAMDENSEKKRKRAIKLIVTLYMLI----- 299
QY 318 SIVASKIHWPTPYLKFAILRALFIPFFFCHYRQTRAYPVFFESTDIFVIGGIAMSF 377
Db 300 CFIPSNL-----LLVYHYFLKSQGSHVYALYIVACLSTLNCSIDPF 343
QY 378 SHGYSALAMGTTPNVPVPSHYSRFAAQLSV 407
Db 344 VYFVSHDERDAHNALLCRSRTVKQMOV 373

RESULT 14
US-08-472-840-6
Sequence 6, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELLIN, JOHAN
APPLICANT: SCARBROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE: 25-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006-20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INSTRUCTION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-840-6

Query Match 4.7% Score 108.5; DB 1; Length 398;
Best Local Similarity 19.2%; Pred. No. 0_0083;
Matches 75; Conservative 59; Mismatches 141; Indels 115; Gaps 18;
OY 61 VETWYS-KEFMGSLTIGSQ---LPNASINVNPNUFLITAGPLIYRVA-----PVCFN 108
Db 56 VEIVFSVDEFSASVLTGKTTVFLPIVIVFVGPSNGMAIW-VELFRTRKKHPAVIY 114
OY 61 VETWYS-KEFMGSLTIGSQ---LPNASINVNPNUFLITAGPLIYRVA-----PVCFN 108
Db 56 VEIVFSVDEFSASVLTGKTTVFLPIVIVFVGPSNGMAIW-VELFRTRKKHPAVIY 114
OY 109 IVNUTILILVIVLEPTEDSMSHFWVTLGMATSINFNSGLYENSYVGDDFPHTYIGA 168
Db 115 MANLALADLSVT-----WFLKIHYHGNWNIGEALCNV-----LIG- 154
OY 169 LLTGNNICGLLTVVKIGTYFLNDEP---KLVATVFGISLVI-LLVCAAL-FFITK 222
Db 155 FFIGNMYCISLFLMTCLSVORYWIVNPMGHSKKANIAIGISLAIWLLVITPLYVK 214
OY 223 QDFYHYHHQKGMEIREKAETDRPSPSLWTFLTFINCY-----GOLFNYWFCFAVLT 273
Db 215 QTIF-----IPAL---NITTCHDVLPEOVLVGDMENYFLSLAIGVF 252
OY 274 IEPVMMTYTRGDSGF-----NKIMSENDE-----IYTLLTSFLVNLFAAIG 317
Db 253 LFP AFLTA-----SAYVLMIRMRSSAMDENSEKKRKRAIKLIVTLYMLI----- 299
OY 318 SIVASKIHWPTPYLKFAILRALFIPFFFCHYRQTRAYPVFFESTDIFVIGGIAMSF 377
Db 300 CFIPSNL-----LLVYHYFLKSQGSHVYALYIVACLSTLNCSIDPF 343
OY 378 SHGYSALAMGTTPNVPVPSHYSRFAAQLSV 407
Db 344 VYFVSHDERDAHNALLCRSRTVKQMOV 373

RESULT 14
US-08-472-840-6
Sequence 6, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELLIN, JOHAN
APPLICANT: SCARBROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

RESULT 15
US-08-476-976-6

```

; Sequence 6, Application US/08476976
; Patent No. 584400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS,
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-476-976-6

```

```

Db 344 VYYFVSHDFRDHAKNALLCRSVRTVKMQV 373
Search completed: February 27, 2002, 16:59:53
Job time: 205 sec

Db 344 VYYFVSHDFRDHAKNALLCRSVRTVKMQV 373
; Query Match 4.78; Score 108.5; DB 2; Length 398;
; Best Local Similarity 19.28; Pred. No. 0083; Indels 115; Gaps 18;
; Matches 75; Conservative 59; Mismatches 141;
Qy 61 VETWYS-KEFGNSLTIGSQ---LPNASINFLTLIAGPLIYRVE-----PVCFN 108
Db 56 VEIVFSDEFASAVLTKLTVFLPVIYIFVVGLPSNGMALW-VELFRKKKHPAVY 114
Qy 109 IVNLTIILIVLVELEPTEDSMSWFVWVTLGMATSNFSNGLYENSYVGCGDFPHTYIGA 168
Db 115 MANLADLISVT-----WEPPLKAYHIHGNNWVGEALCNV-----LIG- 154
Qy 169 LLIGNNCGLITVVKIGTVYFLINDEP---KLVAVYFGSLVILLYCNAIAL-FFITK 222
Db 155 FFYGNMYSCLFMTCLSVQRWVIVNPMGHESRKKANTAIQSLTWLILVTPILYVVK 214
Qy 223 QDPFYHYHQKGMEIREKAETDRPSPLWTFNTCY-----GQLFNWFCFAVLT 273
Db 215 QTIF-----IPAL---NITTCHDVLPLPQVLYGDMFNFLSAGVF 252
Qy 274 IFPVMMTVTTRGDSGFL-----NKIMSENDE-----YTLLTSLFVNLAIG 317
Db 253 LFPAFLTA-----SAYLMIRLSSAMDENSEKKRKRAIKLVTLAMYLI----- 299
Qy 318 SIVASKIHWPTPRYLKAIIIRALFIPFFFCNYRQTRAVPVFFRSSTDIVIGGIAMSF 377
Db 300 CPISNL-----LLVYHFJLKSGOSGHVVALYIVALCSTLNSCIDPF 343
Qy 378 SHGYSALAMGYTPNVVPSHYSRFAAQLSV 407

```

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match Length	DB ID	Description
1	2290	99.8	434	CORK potassium
2	372.5	16.2	475	Rat ENT1 receptor
3	343.5	15.0	456	Human ENT1 receptor
4	343.5	15.0	456	Human ENT1 receptor
5	327.5	14.3	475	Human ENT1 receptor
6	326.5	14.2	456	Equilibrium n
7	326.5	14.2	456	Human equilibrium
8	326.5	14.2	456	Human equilibrium
9	326.5	14.2	475	Human NBPR-1EN
10	326.5	14.2	475	Human PRO1380 C
11	320.5	14.0	456	Rat equilibri

No. pred. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

COMMOTES

result No.	Score	Query Match	Length	DB	ID
1	2230	99.8	434	17	AAV
2	372.5	16.2	475	21	AAV
3	343.5	15.0	456	19	AAV
4	343.5	15.0	456	21	AAV
5	327.5	14.3	475	21	AAV
6	326.5	14.2	456	19	AAV
7	326.5	14.2	456	19	AAV
8	326.5	14.2	456	21	AAV
9	326.5	14.2	475	21	AAV
10	326.5	14.0	456	19	AAV
11	320.5	14.0	456	19	AAV

QY	160 DEPHTYGALLIGNNIGLILVYRIGVTVFLNDPKLVAVYFGLTSLVLYALIFF	219
DDB	192 sfpnnagalisomggtvsvasaki-vdlaassvrdsalafftaavflgicvgly11	250
QY	220 ITKDFYHHOKGMEIREKAETDRP--SPSI-----LWTTFTNCYQOLF	262
DDB	251 lqplearyymrppvphvffssppdrpstssvaprasrvahptplqlpilktagf	310
QY	263 NVNFCFAVLTIFPVMMTTRGDGFLNKIMSENDEIYLTLISFLVNLFAAIGSIVAS	322
DDB	311 cavifitalipatistiqpmhkqspwts-kfypllvflnfadlcgrqvia	367
QY	323 KIHWPPTR--YLKFAILRLALPFFENPYYRVTOPAYPFESTDIF-VIGGIAMSFSH	379
DDB	368 wiqvpqpskllpilavsvclplflcnqyqsh1tvlvfqs-diypilftclglsln	426
QY	380 GYLSLAMCYTPNVYPSHYSRFAAQLSYCTLMGLTGGLWPVYEHFY	428
DDB	427 gyistlvimygpkvprealaatsvmlfymsigmlgsacaailehfi	475
RESULT		
	3	
ID	AAN69555 Standard; Protein: 456 AA.	
XX		
AC	AAN69555;	
XX		
DT	13-OCT-1998 (first entry)	
DE	Human equilibrative nucleoside transporter 1.	
KW	Human; equilibrative nucleoside transporter; hENT1; hENT2; rENT1; rENT2; coronary; cerebrovascular anoxia; viral infection; cancer.	
KW		
XX		
OS	Homo sapiens.	
XX		
PN	WO9829437-A2.	
XX		
PD	09-JUL-1998.	
XX		
PP	30-DEC-1997; 97W0-1B01657.	
XX		
PR	03-NOV-1997; 970S-0064.004.	
PR	30-DEC-1996; 96US-0034083.	
XX		
PA	(UYAL-) UNIV ALBERTA.	
PA	(UYLE-) UNIV LEEDS.	
XX		
PI	Baldwin SA, Cass CE, Young JD;	
XX		
DR	WPI; 1998-388035/33.	
DR	N-PSDB; AAV40276.	
PT	Newly isolated equilibrative nucleoside transporter protein(s) and gene(s) - used to develop products for treating disorder(s) associated with the transporter(s) and for use with nucleoside drug(s)	
XX		
PS	Claim 5; Fig 3; 97pp; English.	
XX		
CC	The present sequence represents a substantially purified equilibrative nucleoside transporter (ENT), human ENT1 (hENT1). ENTs can transport a variety of purines and pyrimidines, including adenosine, uridine, guanosine, inosine, formycin B, tubercidin, and thymidine. ENTs are bidirectional, they can be used as a tool for the development of new nucleosides drugs. Products from the present invention can be used for treating a subject having a disorder associated with an ENT. They can also be used with nucleoside drugs in the treatment of e.g. coronary or cerebrovascular anoxia, viral infection or cancer. The products (e.g. antibodies and oligonucleotides hybridizing to nucleic acid sequences encoding ENTs) can also be used for detection and drug screening.	
CC		

PT therapy -
 XX Disclosure: Fig 2A-B; 55PP: English.
 PS
 XX The invention relates to the isolation of a gene encoding a NBMPR-iENTP (nitrobenzylmercaptoxypurine-riboside-insensitive, equilibrative nucleoside transporter protein) from HeLa cells. The protein transports nucleosides across the plasma membrane by a facilitated diffusion process and is involved in the proliferative response. The transporter is useful in screening to identify natural nucleoside permeants and/or their inhibitors or analogues, potential therapeutic agents, also for studies on protein structure and mechanism. Cells that express the iENTP as the only transporter protein are used for drug screening (especially to identify antitumour and antiviral nucleoside analogues), in chemotherapy of cancer and for selective expression of heterologous genes for gene therapy. The cDNA is used for recombinant expression of iENTP and as a source of oligonucleotides (diagnostic primers and probes, ribozymes and antisense sequences). Antibodies raised against iENTP are used for detection of the protein by usual immunoassays and as (antagonists of iENTP activity. This sequence represents the human ENTR1 protein (AAI5517).

XX Sequence 456 AA:
 SQ

Query	Match	Score	Length	DB	No.	Pred.	Indels	Gaps	Best Local Similarity	Matches	Database	Conservative	Deletions	Insertions	Amino Acids
Qy	18	PRDKNIVYWVLVILVGFGVLLPWNMFITIAPEYYVNVWFKPDGVE--TWYSKEBFMGSLLT	343.5	21	456;				24.3%	112;	Human NBMPR-iENTP				
Db	7	pqrkykavllffmglgtlpwnffmt-atqyfnrlamsqnslvtaelskdqasaa	74												
Qy	75	IGSQLP--NASINVNENLFLITIAGPLIYRFA-----PVCFENIYN----LTITLI	117												
Db	66	paappperisaiinnyntcamplifllylnsfhlgrqsvrlqslvavillfl	125												
Qy	118	LIVILEPTEDMSMNFWWFTLGMATSINFNSGLYENSGLVGGDFPHTYIGALLIGNNICG	177												
Db	126	tailvkqvqidalp-ffvitmkivinstgailqslglagllpasytapimsgqqiag	184												
Qy	178	LITIVVKICIVYFLNDEPKLVAIVYFGISLVLVLCATALFVTKDFHYHHQ-----	231												
Db	185	ffasvami-caiasqeiseafayfitacavillticylgiprfryqqqlklegp	243												
Qy	232	-KGNEIREXATAEDTPSPSITWTT-----PINCYGOLENWVFCFAVTT	271												
Db	244	gegetkldiskgeprakteseyvsvisqptneshskiaikknisvlafscfcifitit	303												
Qy	272	LTIFPVMMVTTRGSGFLNKIMSENDEIYTLLTSELYFNLEAIGSIVASKIHWP--TP	329												
Db	304	igmpavtvekssqtag----sstwetyfipyscfltinfidwlgrsitasvwmpgkds	358												
Qy	330	RYLEFAIIILRALFIPFFFONRYVOTRAY-PVFFESTDIFVIGGIAMSFSGYLSALANG	388												
Db	359	rwlpsvlarlvfpflilcn--ikprryltvvfehdawffmafafsnqyaslcmc	416												
Qy	389	YTPNPVPSHYSRFAQLSVCLMYGLPWPVIEHV	428												
Db	417	f9pkkvkpaetaetgaimaffciglalgavstflfraiv	456												

RESULT 5
 AAYB2285 standard; Protein: 475 AA.
 ID AAYB2285
 XX
 AC AAYB2285;
 XX
 DT 19-JUN-2000 (first entry)
 DE Human ENTR1 receptor SEQ ID NO:1.
 XX Human; ENTR1; immunosuppressant; vasotropic; thrombolytic;
 KW

Query Match Score 327.5; DB 21; Length 475;
 Best Local Similarity 24.4%; Pred. No. 3.3e-25;
 Matches 116; Conservative 83; Mismatches 202; Indels 75; Gaps 15;

Qy 7 SNTYAVEOAF-----PRDYNIVYWLVILVGFGVLLPWNMFITIAPEYYV 52
 Db 21 sss1radallekkldlppglqrpedrcgylifsglslpwwffit-akewwm 79

Qy 53 NYFWK-----PDGVETWYSKEEMGSLTIGSQLPNASINVNFNUFLIIAGPLIYR 100
 Db 80 ---fklnrissspagedpgsd---ilnlyfesylavastvpsmlcvnlvnvavhr 134

Qy 101 VFAPVCFNIVNLTLILIVLVEP-TEDMSMW--FWFWTLMGATSNFSNLGYNESVY 155
 Db 135 vla-----sltvvlairmvitalvkvdtrgffavtivcmvisgastvfvsssy 187

Qy 156 GVGDGFPHYIGALLIGNNCGLLIVYFLNDPPLKLVLLVLLYCAI 215
 Db 188 gmcgsfmrsqalsqgmggqgvssavsi-vdlaassdvrsnalsaffatifiving 246

Qy 216 ALFFITKQDFHY-----HIIOKGMREKAEIDRSPSPSI-----LWTTEFTN 256
 Db 247 lyllsrleyarymrpvaahvsgee--elpgdslapsvsrfishhtpprpikk 304

Qy 257 CYGQLFNWVFCCFAVLTIEFPVMMTVTRGDSGLEINKINSE--ADEIYLTTSFLVFNLF 314

DDB	305	tas1gfcvtfvffitlliypavt---nieslnkgsgslwttkffipittflynfad	359
QY	315	AIGSIVASKIHWPTP--RYLKFAILRLAFLIPFFFCNRYQVTRAYPVFFFESTDIRWIGG	372
DDB	360	1cgqrqitawlqvpgbnkskaalpgfvilrtcliplfvcnysqprvhktvvfjsdvpypalls	419
QY	373	IAMSFSHGVLUSALANGYTPNVPVSHYSRFAAQQLSNVCTLMGLLTGGILWPVVIEHFV	428
DDB	420	sliglsngylstlalliygpkvprelaeatgvmsfsyvcigltsgascstlvhlri	475
RESULT 6			
	AAW83965	AAW83965 standard; Protein; 456 AA.	
	XX		
	AC		
	XX		
	DT		
	DE	Equilibraive nucleoside transport protein (IENTP).	
	XX		
	FT		
	Domain	Equilibraive nucleoside transport protein; IENTP; NBMPR; transport;	
	FT	nitroimidazole; antitumour; antiviral; antitumour; screening;	
	FT	inhibitor; drug; knockout mice; cancer; chemotherapy; gene therapy;	
	FT	adenosine deaminase; factor VIII.	
	XX		
	OS	Homo sapiens	
	XX		
	Key	Location/Qualifiers	
	FH	13..27	
	FT	/note= "transmembrane domain TM1"	
	FT	71..93	
	FT	/note= "transmembrane domain TM2"	
	FT	100..117	
	FT	/note= "transmembrane domain TM3"	
	FT	125..144	
	FT	/note= "transmembrane domain TM4"	
	FT	162..185	
	FT	/note= "transmembrane domain TM5"	
	FT	194..214	
	FT	/note= "transmembrane domain TM6"	
	FT	293..312	
	FT	/note= "transmembrane domain TM7"	
	FT	326..341	
	FT	/note= "transmembrane domain TM8"	
	FT	361..379	
	FT	/note= "transmembrane domain TM9"	
	FT	393..413	
	FT	/note= "transmembrane domain TM10"	
	FT	432..452	
	FT	/note= "transmembrane domain TM11"	
	XX		
	PN	WO9846749-A1.	
	XX		
	PD	22-OCT-1998.	
	XX		
	PF	98WO-US07283.	
	PF	10-APR-1998;	
	XX	98US-0058389.	
	PR	09-APR-1998;	
	PR	97US-0838845.	
	PR	11-APR-1997;	
	XX	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
	PA		
	XX		
	P1	Belt JA, Crawford CR, Patel DH;	
	XX		
	DR	WPI; 1998-594576/50.	
	DR	N-PSDB; AAV69638.	
	XX		
	PT	New equilibraive nucleoside transport protein insensitive to	
	PT	nitrobenzylthio-purine riboside - useful for, e.g. identifying to	
	PT	specific therapeutic nucleoside analogues and in gene therapy to	
	PT	protect transduced cells against ablative chemotherapy	
	XX		

QY	376 SFSGCGYL SALAMGTYTPNVVPSHYSRFAAQLSVCTLMVGLLTG 417
Db	404 avsnqylvsutmc1qrq1pherevagalm1mtff1alq1scg 445
RESULT 8	
AAB15517	
ID AAB15517	standard; protein: 456 AA.
XX	
AC AAB15517;	
XX	
DT 14-FEB-2001	(first entry)
XX	
DE Human NBMPR-iENTP protein.	
XX	
KW Antiviral; antitumour; NBMPR-iENTP; nitrobenzylmercaptopurineriboside;	
KW insensitive, equilibrative nucleoside transporter protein; cancer;	
KW plasma membrane; facilitated diffusion; gene therapy; primer; probe;	
XX	
OS Homo sapiens.	
PN US6130065-A.	
XX	
PD 10-OCT-2000.	
XX	
PF 09-APR-1998;	98US-0058389.
XX	
PR 11-APR-1997;	97US-0043659.
XX	
PA (SJUD) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	
PI Crawford CR, Belt JA;	
XX	
DR WPI: 2000-637839/61.	
DR DR-N-PSDB; AAA95722.	
XX	
PT New nucleic acid encoding nucleoside transporter, useful for drug	
PT screening to identify antitumour and antiviral agents and for gene	
PT therapy -	
XX	
PS Claim 1; Fig 2A-B; 55pp; English.	
XX	
The present sequence represents a substantially purified equilibrative nucleoside transporter (ENT), human ENT2 (hENT2). ENTs can transport a variety of purines and pyrimidines, including adenosine, uridine, guanosine, inosine, formycin B, tubercidin, and thymidine. ENTs are bidirectional, they transport a suitable permeant both into and out of cells. ENTs can be used as a tool for the development of new nucleoside drugs. Products from the present invention can be used for treating a subject having a disorder associated with an ENT. They can also be used with nucleoside drugs in the treatment of e.g. coronary or cerebrovascular anoxia, viral infection or cancer. The products (e.g. antibodies and oligonucleotides hybridising to nucleic acid sequences encoding ENTs) can also be used for detection and drug screening.	
XX	
Sequence 456 AA:	
Query Match 14.28; Score 326.5; DB 19; Length 456;	
Best Local Similarity 22.9%; Pred. No. 4e-25;	
Matches 106; Conservative 71; Mismatches 200; Indels 85; Gaps 11;	
Dy 18 PRDKYNIVWLYLIVGRGVLLPWNMFITIAPYY-----VNWFKPDGVE 62	
Dy 7 Pdyslhvgisfflglgtlpwnffitaiyfgarlagagnstarllsthtngpedaf- 65	
Dy 63 TWYSKEFMGSTIGSQLPNASINVNFLLINGPLIYRFAPVCENVNLTILIVVL 122	
Dy 66 -----nfnnwttllsqplllfltsflycypetvrl9---sillaillfaltaal 116	
Dy 123 EPTEDMSWFWVNTLGMASTINFNGLYENSYVGVDGFPTIYIGALLIGNNCGLLTV 182	
Dy 117 vkvdmsspgffitmasvcifnsavlgqslfqgqtmpstytf1sggglagifaal 176	
Dy 183 VKI-----GTYFLNDEPKLVIAIVF---GTSILVLUCA1ALFFTQDFY---HYHHQ 231	
Dy 177 amilsnasgv-----daetsalgylfitpyrglmsivcysiphikfarylankssqa 230	
Dy 232 KGMEIREKA-----ETDRPSPLSLWTTFTNC 257	
Dy 231 qaelekael1qsdengipspqkvaltlldiekepessedepakpkpsvftfqki 290	
Dy 258 YGQLFNNWFCAVLTIFPVMMTIVTRGDSGLNKIMSENDEYTLLTFSVLFNLEAIG 317	
Dy 291 vitalcivlvttvtlsvpfaatmsts-----pgkwqsfnnpicclflnlmdwlg 344	
Dy 318 SIVASK1HWP---TPYRLKFAITLRAFLPFEFFFCNYVRQTVAYPFFESTIDFVGGIAM 375	
Dy 345 rsitsyflwpdedsrllpllvclrlfvplmlchvqqrst-1p1lpqddayfitfmllf 403	
Query Match 14.28; Score 326.5; DB 21; Length 456;	
Best Local Similarity 22.9%; Pred. No. 4e-25;	
Matches 106; Conservative 71; Mismatches 200; Indels 85; Gaps 18 PRDKYNIVWLYLIVGRGVLLPWNMFITIAPYY-----VNWFKPDGVE 62	
Db 7 prsyhlgisfflglgtlpwnffitaiyfgarlagagnstarllsthtngpedaf- 65	
Qy 63 TWYSKEFMGSTIGSQLPNASINVNFLLINGPLIYRFAPVCENVNLTILIVVL 122	
Db 66 -----nfnnwttllsqplllfltsflycypetvrl9---sillaillfaltaal 116	

QY 123 EPTEDSMSWFWTIGMATSINFNSGLYENSYVGDDPFPHTYIGALLIGNICGLITY 182
 DB 117 Vkydmsppgpfisnasytisavsgisifgg1q1unpststflsqgqlagfaal 176
 QY 183 VKI----GVTYFLDEPKLVAYVFE--GISLVLVLLCAIALPEPITKODFY--HYHHIQ 231
 DB 177 amlsmsgy-----daertsalyfitpcvgimsvicylslphkfarylankssqa 230
 QY 232 KGMETREKA-----ENDRSPSPSLWTTFTNC 257
 DB 231 qaqeletkaellqsdengipspqkvvaltdidlekepesepdepqkpkpsvttvfqki 290
 QY 258 YGQLFNWNFCFAVLTIFPVMMTYTRGDSGFNLKIMSENDEIYTLLTSFLVENLFAAIG 317
 DB 291 witaclcvlvitvtttsvfpaitanvtssts----pqjwsqfninpccilfrimdwq 344
 QY 318 SIVASKIHW--TPRYLKFAITLRAFLIPPFCCNRYVQTRAYPVFFESTDIFVIGGIAM 375
 DB 345 rsllsytflwpdedsr1pllvclrlfvlpfm1chvpqrsr-1p1lfpqdayfitfml1f 403
 QY 376 SFSHGYSALAMGXTPNVVPSHYSRFAAQLSVCTLMVGLLTG 417
 DB 404 avsngylvsitmc1aprqvlpherevagalmtffialgscg 445

RESULT 9
 ID AAY9363 standard; Protein: 475 AA.
 XX AAY9363;
 AC XX
 DT 08-AUG-2000 (first entry)
 XX DE Human PRO1380 (UNQ717) amino acid sequence SEQ ID NO:79.
 KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
 KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
 OS Homo sapiens.
 XX FN WO200012708-A2.
 XX PD 09-MAR-2000.
 XX PF 01-SEP-1990; 99WO-US201111.
 XX PR 01-SEP-1998; 98US-00987116.
 PR 01-SEP-1998; 98US-0098749.
 PR 01-SEP-1998; 98US-0098750.
 PR 02-SEP-1998; 98US-0098803.
 PR 02-SEP-1998; 98US-0098821.
 PR 02-SEP-1998; 98US-0098843.
 PR 09-SEP-1998; 98US-0098843.
 PR 09-SEP-1998; 98US-0099336.
 PR 09-SEP-1998; 98US-0099596.
 PR 10-SEP-1998; 98US-0099580.
 PR 09-SEP-1998; 98US-0099588.
 PR 09-SEP-1998; 98US-0099602.
 PR 09-SEP-1998; 98US-0099612.
 PR 10-SEP-1998; 98US-0099711.
 PR 10-SEP-1998; 98US-0099815.
 PR 10-SEP-1998; 98US-0099815.
 PR 15-SEP-1998; 98US-0099815.
 PR 15-SEP-1998; 98US-0099815.
 PR 15-SEP-1998; 98US-0100388.
 PR 15-SEP-1998; 98US-0100388.
 PR 16-SEP-1998; 98US-0100390.
 PR 16-SEP-1998; 98US-0100394.
 PR 16-SEP-1998; 98US-0100667.
 PR 16-SEP-1998; 98US-0100667.
 PR 16-SEP-1998; 98US-0100664.
 PR 03-NOV-1998; 98US-0106856.
 PR 03-NOV-1998; 98US-0106902.

PR	03-NOV-1998;	98US-0106905.	Qy	257 CYGOLFNYWFCFAVLTIFPVMMTVTTRGDSGFNLNKIMSE--NDEIVTLLTSELVNFNA 314
PR	03-NOV-1998;	98US-0106919.	Db	305 taslgfcyyffeltlypavct---niesinksgslwtktffplttfllynfad 359
PR	03-NOV-1998;	98US-0106932.	Qy	315 AIGSTIVASSXIHWPDP--FYLKFAILRFLIPPFECMYRVQTRAYPPFEESTDIFVFGG 372
PR	03-NOV-1998;	98US-0107783.	Db	360 1cgqrqltawiqveognshalpgflrlrciplfvlnqprvhltkvfqsdvypal s 419
PR	17-NOV-1998;	98US-0108775.	Qy	373 IAMSFSHGYLSALAMGYTPNVVPSHYSRFAAQSLVCUMLMVGLJTGGLWPMVIEHFV 428
PR	17-NOV-1998;	98US-0108779.	Db	420 slqisnqytlialyqpkivprelaatgvvmsfyvcigltqsaestlvhl i 475
PR	17-NOV-1998;	98US-0108787.		
PR	17-NOV-1998;	98US-0108801.		
PR	17-NOV-1998;	98US-0108802.		
PR	17-NOV-1998;	98US-0108806.		
PR	17-NOV-1998;	98US-0108807.		
PR	17-NOV-1998;	98US-0108807.		
PR	17-NOV-1998;	98US-0108825.		
PR	18-NOV-1998;	98US-0108848.		
PR	18-NOV-1998;	98US-0108849.		
PR	18-NOV-1998;	98US-0108850.		
PR	18-NOV-1998;	98US-0108851.		
PR	18-NOV-1998;	98US-0108852.		
PR	18-NOV-1998;	98US-0108858.		
PR	18-NOV-1998;	98US-0108904.		
XX				
PA (GETH) GENENTECH INC.				
PI Baker K,	Goddard A,	Gurney AL,	Smith V,	Watkanabe CK,
XX				Wood WI;
DR	2000-237871/20.			
XX				
PT N-FSDB; AAA37045.				
PT New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides, useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions				
PT				
PT				
XX				
PS	Claim 12; Fig 48; 773pp; English.			
XX				
CC	AAA37022 to AAA37144 encode the new isolated human transmembrane, receptor or secreted PRO polypeptides given in AAY93340 to AA99462. The transmembrane and receptor PRO proteins can be used for screening of			
CC	potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoding them have various industrial applications, including uses as pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR primers and hybridization probes used in the isolation of the PRO polypeptides from the present invention.			
XX				
Sequence	475 AA;			
Score	326.5;	DB 21;	Length 475;	
Best Local Similarity	24.48;	Pred. No. 4.2e-25;	Indels 75; Gaps 15;	
Matches	116; Conservative			
Qy	7 SNTYAVEQBAF-----PRDKYNVYWLVLGVLLPNMFITIAPPYY 52			
Db	80 --fklnrsspatgedpegsd -ilnyfesyyavastvpmsnlcyanflivnravhir 134			
Qy	101 VFAAPVCFENIVNNTLILVILEP--TEDMSW--FFWVTLGMATSNFSNGLYENSY 155			
Db	135 via -----stivalifmvlavkvdtswtrgffavtvcmlvilsagtfssiy 187			
Qy	156 GWGDFPHTYIGALLIGNNCIGLITVKIGSTYFLNDEPKLVIAIVFGISIVILLVCAI 215			
Db	188 gmtgsfpnrrsalsqgammgqsvavasi -rliaasdvrsnraafflatflvcmg 246			
Qy	216 ALFIFTKDDFYH-----RHQKGMEIREKAETDRPSPSL-----LWTTFTN 256			
Db	247 lyllsrieyarymrpvlaanhfsqee--epqdsisapsrasfdhtpprlplkk 304			
Query Match	14.28;	Score 326.5;	DB 22;	Length 475;
Best Local Similarity	24.48;	Pred. No. 4.2e-25;	Indels 75; Gaps 15;	
Matches	116; Conservative			

RESULT 11
AAW69557
ID AAW69557 standard: Protein: 456 AA

AAW65557;			
XX			
DT	13-OCT-1998	(first entry)	
XX			
DE	Rat equilibrative nucleoside transporter 2.		
XX			
KW	Rat; equilibrative nucleoside transporter; hENT1; hENT2; rENT1; rENT2; coronary; cerebrovascular anoxia; viral infection; cancer.		
XX			
OS	Rattus sp.		
XX			
PN	W0989437-A2.		
XX			
PD	09-JUL-1998.		
XX			
PF	30-DEC-1997;	97WO-IB01657.	
XX			
PR	03-NOV-1997;	97US-006404.	
PR	30-DEC-1998;	96US-0034083.	
XX			
PA	(UYAL-) UNIV ALBERTA.		
PA	(UYLE-) UNIV LEEDS.		
XX			
PI	Baldwin SA, Cass CE, Young JD;		
XX			
DR	WPI; 1998-388035/33.		
DR	N-PSDB; AAV40278.		
XX			
PT	Newly isolated equilibrative nucleoside transporter protein(s) and gene(s) - used to develop products for treating disorder(s) associated with the transporter(s) and for use with nucleoside drugs(s)		
XX			
PS	Claim 5; Fig 14; 97pp; English.		
XX			
Db	237 ae1lgadekngipspqqagptldldpekelelgiepapkpgkpsvfrkiwltacl 296		
QY	265 WFCFAYVLTIFP-WMMTVTRGDSFLNKIMSENDEIYLTLTSLVFLFAIGSIVASK 323		
Db	297 vlvttvlsvfpalitamvtssns ----pgkwsqffpicccflfnmdwgrslyes 350		
QY	324 IHWP -TPRLKFALLRALFIPPEFCNVRQTRAPYFFESTDIFVGGIAMSFSHGX 381		
Db	351 flwpdedsqlpl1vcrlrifvlpmlchqvqrar-lp1ifwgdayfitfmlfaisnsg 409		
QY	382 LSA1AMGYTPNVVPSHYSRFAAQLSVCTLMVGILTG 417		
Db	410 fvsitmc1aprqvhvarevagaintfflqlscg 445		
		RESULT 12	
		AAW65556	
		AAW65556 standard; Protein; 457 AA.	
		ID	
		AAW65556;	
		AC	
		AAW69556;	
		XX	
		DT	
		13-OCT-1998 (first entry)	
		XX	
		DE	
		Rat equilibrative nucleoside transporter 1.	
		XX	
		KW	
		rENT2; coronary; cerebrovascular anoxia; viral infection; cancer.	
		XX	
		OS	
		Rattus sp.	
		XX	
		PN	
		W09829437-A2.	
		XX	
		PD	
		09-JUL-1998.	
		XX	
		PF	
		30-DEC-1997;	

Db 291 111en--ikprryltvvfehdawffffmaafafsngylas1cmcfcgpkvvpaaetaetg 348
 Qy 404 QLSVCTLMGYGLTGGLMPVYIHFV 428
 Db 349 aimafflclgalgavfsffirai 373

RESULT 14
 PAB93550
 ID AAB93550 standard; Protein: 397 AA.
 XX
 AAB93550;
 AC
 DT 26-JUN-2001 (first entry)
 XX
 Human protein sequence SEQ ID NO:12928.
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FN EP1074617-A2.
 XX
 PD 07-FEBR-2001.
 XX
 FF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1998; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183776.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.

XX
 PR primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -
 XX
 PS Claim 8: SEQ ID 12928: 2537pp + CD ROM: English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 3'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH1363 to AAH18742 represent human cDNA sequences; AAB92446 to AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.
 XX
 Sequence 397 AA;

Query Match 11.4%; Score 261.5; DB 22; Length 397;
 Best Local Similarity 23.7%; Pred. No. 1.7e-18;
 Matches 92; Conservative 72; Mismatches 181; Index 43; Gaps 10;
 Qy 69 FMGSILTSOLPNAISNFNLFLIIAGPLIYRVFAPCFNIVNLTLILIVLVEP--TE 126
 Db 25 fesylavasavpsmclvanflvnrvhvirvla ---silviafmvitalvkv 77.
 Qy 127 DMSW--FFWVTLGMATSINFNSGLYENSYVGCGDFPHTYIGALIGNNNICGILITVV 183
 Db 78 dtsswrgflavtivcmvisgasftvssiygmgsfpmrnsqaisggmaggvsava 137
 Qy 184 KIGVYFLNDEPKLYAVVYRGGISLVLYLIVCAIALEFFITKQDFEHY -----HRQKME 235
 Db 138 sl-vdlaassovrnsalaffitflvlenglyslslearyympvlaahvfsgsee 196
 Qy 236 IREKAETDRPSPT-----LWTTFTNCYGOLENWFCFAVLTIFPVMMTVTTR 284
 Db 197 --elpodslsapsvsarfidshtpprlplkktaslgfcvtyffitstlypavct--- 250
 Qy 285 GDSGLFLNKINSE--NDEIYIUTLTSLEVNLFAAIGSIVASKLHWPTP--RYLKFAAILRA 340
 Db 251 -niesinksgslwttkffipltfllynadlcgrqtawiryppgnshalpgfvilrt 309
 Qy 341 LFIPFFFCNYRQVQTARYPAFFESTDIFVIGGIAMSFSHGYSALAMGYTPNVPDSHYSR 400
 Db 310 cliplylcnqyprvhktvqfqsadvpalissllsgnysitallaypkivprellae 369
 Qy 401 FAAQLSVCTLMVGLTGJGLWPVVIIEFH 428
 Db 370 atgvvmsfywlgltqsgactslvhl 397

RESULT 15
 AAB15519
 ID AAB15519 standard; Protein: 326 AA.
 XX
 AC AAB15519;
 XX
 DT 14-FEB-2001 (first entry)
 DE Human HNP36 protein.
 XX
 KW Antiviral; tumour; NBMPR-1ENTP; nitrobenzylmercaptapurinoriboside; KW insensitive, equilibrative nucleoside transporter protein; cancer; KW plasma membrane; facilitated diffusion; gene therapy; primer; probe.
 XX
 OS Homo sapiens.
 XX
 US6130065-A.
 XX
 PN 10-OCT-2000.
 XX
 PR 09-APR-1998; 98US-0058389.
 XX
 PR 11-APR-1997; 97US-0043659.
 XX
 PA (SJUD) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 PI Crawford CR, Belt JA;
 XX
 DR WPI: 2000-637839/61.
 XX
 PT New nucleic acid encoding nucleoside transporter, useful for drug screening to identify antitumour and antiviral agents and for gene therapy -
 XX
 PS Disclosure; Fig 2A-B; 55pp; English.
 XX
 CC The invention relates to the isolation of a gene encoding a NBMPR-1ENTP
 CC (nitrobenzylmercaptapurinoriboside-insensitive, equilibrative nucleoside
 CC transporter protein) from HeLa cells. The protein transports nucleosides
 CC across the plasma membrane by a facilitated diffusion process and is

CC involved in the proliferative response. The transporter is useful in
 CC screening to identify natural nucleoside permeants and/or their
 CC inhibitors or analogues, potential therapeutic agents, also for studies
 CC on protein structure and mechanism. Cells that express the iENTP as the
 CC only transporter protein are used for drug screening (especially to
 CC identify antitumour and antiviral nucleoside analogues), in chemotherapy
 CC of cancer and for selective expression of heterologous genes for gene
 CC therapy. The cDNA is used for recombinant expression of iENTP and as
 CC a source of oligonucleotides (diagnostic primers and probes, ribozymes
 CC and antisense sequences). Antibodies raised against iENTP are used for
 CC detection of the protein by usual immunoassays and as (anti)agonists of
 CC iENTP activity. This sequence represents the human HNP36 protein and
 CC is used for comparison with the human NBMPR-iENTP protein (AAB15517).

SQ Sequence 326 AA:

	Query Match	Best Local Similarity	Score	DB	Length
Qy	143 INFENGLYENSYVGGVGGDFPHTYIGALLIGNICGLLITVVKI-----	214;	DB 21;	Length 326;	
Db	7 insfsaviqsq1qsfqg1qmpstiflsqqqagfaalamilsmasgv-----daet	52;	Pred No 1e-13;	Mismatches 139;	Indels 60; Gaps 8;
Qy	198 VAIYF--GISLYVILLVCAIALEFFITQDFY--HYHHHQGMETIREKA-----	214;	DB 21;	Length 326;	
Db	61 saqyfitpyvgilmssvrylankssqaqcletkaellqsdenqip	120;			
Qy	241 -----ETDRBPSLILWTTFNCYGOLFNYWFCFAVTLTIFPV	277;	DB 21;	Length 326;	
Db	121 ssdqkvaltldidlekepepepedpqkpgpsvtfvqkiwitalclvlfvtlsvfpa	180;			
Qy	278 MMIVYTTRGDGFNLNKIMSENDEIYLTLTSFLVNLFAIGSIVASKLHWPP-TPRYLKFA	335;	DB 21;	Length 326;	
Db	181 itanwvtssts-----pkwsgqfinpicoflnimdwigrsitsyflwpddesrlpli	234;			
Qy	336 IILRALFIPFFFNCYRQTRAPVFESTDIEVGGTAMSFSHGILSALANGYTPNVP	395;	DB 21;	Length 326;	
Db	235 vciirflfpmlmichvpqrsl-pilfpqdayfitfmilfavaasnqylvslnclaprvql	293;			
Qy	396 SHYSRFAAQQLSVCTLMVGLTG	417;	DB 21;	Length 326;	
Db	294 pherevagalmtfflaiglscg	315;			

Search completed: February 27, 2002, 16:59:32
 Job time: 199 sec